A Competition Mechanism for a Homeotic Neuron Identity Transformation in *Caenorhabditis elegans*

Patricia M. Gordon

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2015

ABSTRACT

A Competition Mechanism for a Homeotic Neuron Identity Transformation in *Caenorhabditis elegans*

Patricia M. Gordon

As embryos proceed through development, they must undergo a series of cell fate decisions. At each division, potency is progressively restricted until a terminally differentiated, postmitotic cell is produced. An important part of that cell type determination is repression of alternative fate possibilities. In this thesis, I have explored the mechanisms by which a single transcription factor activates certain cell fates while inhibiting others, using the *Caenorhabditis elegans* ALM and BDU neurons as a model. ALM neuron identity is regulated by two interacting transcription factors: the POU homeobox gene *unc-86* and the LIM homeobox gene *mec-3*. I investigated fate determination in BDU neurons, the sister cells of ALM. I found that BDU identity is broadly defined by a combination of *unc-86* and the Zn finger transcription factor *pag-3*, while the neuropeptidergic subroutine of BDU is determined by the LIM homeobox gene *ceh-14*. In addition, I found that reciprocal homeotic transformations occur between ALM and BDU neurons upon loss of either *mec-3* or *pag-3*. In *mec-3* mutants, ALM neurons acquire the gene expression profile and morphological characteristics of BDU cells, while in *pag-3* mutants, BDU neurons express genes normally found in ALM and change some aspects of their morphology to resemble ALM. While these fate switches appear to be a simple case of cross-repression, the mechanism is in fact more complicated, as *pag-3* is expressed not just in BDU but also in ALM. In this thesis, I

present evidence that MEC-3 inhibits execution of BDU identity in ALM by physically binding to UNC-86 and sequestering it away from the promoters of BDU genes. This work expands upon the literature examining simultaneous activation of one identity program and repression of alternate programs by introducing a novel mechanism by which a transcription factor competes to direct specific cell fates.

Table of Contents

List of Tables and Figures

CHAPTER 1

CHAPTER 2

APPENDIX A

APPENDIX B

APPENDIX C

Acknowledgements

I am incredibly grateful for the support I've received both during my years in graduate school and in the years leading up to it.

I am extremely fortunate to have been mentored by some amazing scientists. During my senior year of high school, Neil Olszewski at the University of Minnesota first gave me an opportunity to do real science. He and his postdoc Lynn Hartweck gave me an enthusiasm for research that is still with me today. At Vassar College, Kate Susman was both a friend and a mentor. She showed me the beauty and power of *C. elegans* and pushed me to excel.

I cannot say enough about how grateful I am that I've been able to work in Oliver Hobert's lab during my time in grad school. He is an incredible scientist, and it has been an honor learning from him. I also want to thank all the members of the Hobert lab, both past and present, for the support they've given me. Every single person in the lab has been a valuable resource and has given me help, advice, and friendship. In addition, I am grateful to Iva Greenwald and the members of her lab for wonderful scientific discussions at our joint group meetings.

Richard Mann and Wes Grueber have been on my committee throughout grad school, and I am grateful for their questions and advice. I am particularly grateful to Richard, who let me perform EMSAs in his lab, and to his postdoc Namiko Abe, who spent a great deal of her time helping me with those experiments. I also want to thank Marty Chalfie and Max Gottesman, who joined in my thesis committee.

Finally, I am so thankful for the friends and family who have given me love and support over the years. My friends in Minnesota and New York – as well as those

scattered around the world – have kept me sane in times of stress. I particularly want to thank my colleagues at Columbia: Ashlee Bennett, Lindsie Goss, Maria Salee, Kate Stafford, and Holly Wolcott. It was such a pleasure to go through grad school with them.

My family is the greatest family I could ever ask for. My sister Kat Gordon is the best and has been there for me through thick and thin. My in-laws – Paul Vartanian, Joanne Smyth, and Emily Smyth Vartanian – have been wonderful and welcoming since the day I met them. My parents, Louis and Jennie Gordon, have been my biggest cheerleaders throughout my life and have always encouraged me to pursue my dreams. And my husband Adam Vartanian is more loving and supportive than I can describe in words. So many thanks to all of them.

For my family

CHAPTER 1: INTRODUCTION

This introduction will be divided into three main sections. As repression of aberrant cellular programs is intrinsically linked to the correct execution of cell fates, the first section of this introduction will examine the mechanisms of cell fate determination. In particular, I will focus on the establishment of neuronal identity –a complex developmental task that includes both the adoption of general neuronal features and the activation of genes that will allow specific neuronal functions. While it is possible to classify neurons in many ways, one important distinction is their neurotransmitter identity, as the choice of neurotransmitter or neuropeptide drastically alters the functions of a neuron. I will present evidence for the terminal selector theory, showing that in both vertebrates and invertebrates, terminal selectors control cell fate by directly binding to common motifs on the battery of genes that together define the terminally differentiated neuron.

In the second section of this introduction, I will examine the counterpart to cell fate determination: repression of alternate fates. Because failure to inhibit the adoption of incorrect fate programs leads to a homeotic transformation of one cell or tissue into another, I will discuss examples of homeosis in both *Drosophila* and *C. elegans*. I will then look more closely at the mechanisms of alternate fate repression, examining how signaling pathways or transcription factors interact to induce transcription of some genes but not others.

In the third section, I will introduce the system that I used to study these developmental questions: the ALM and BDU neurons in *C. elegans*. This section will examine both the morphology and function of the two neurons and the regulation of

identity in each neuron type. Finally, the last section will present an overview of the thesis.

Part I: Neuronal Development and Cell Fate Determination

1. Neuronal diversity and complexity

Neurons were first described in the 1830s by Christian Gottfried Ehrenberg and Jan Evangelista Purkinje (López-Muñoz et al., 2006), but it wasn't until several decades later that the brilliant work of Santiago Ramón y Cajal revealed the complex and varying structure of the nervous system (Piccolino et al., 1989). Using a variation on Golgi's method of silver staining, Ramón y Cajal showed that neurons constitute discrete cells, rather than a syncytium of anastomoses as had been argued by Golgi (Andres-Barquin, 2002). Ramón y Cajal's legacy lives on not only in the continued triumph of his neuron doctrine (López-Muñoz et al., 2006) but also in his intricate drawings of different neuron types (Pablo Garcia-Lopez, 2010). These drawings, many of which are currently stored at the Cajal Museum in Madrid, show a wide variety of neuron morphologies even within a single cerebral section.

Continued work in neuroscience over the past century has increased our understanding of the incredible number and diversity of neurons in vertebrates. Recent estimates using the isotropic fractionator method have found approximately 90 billion neurons in adult human brains (Azevedo et al., 2009). Each neuron makes between 5,000 and 200,000 synaptic connections with other neurons, leading to the staggering number of 0.15 quadrillion total synapses in a single brain (Muotri and Gage, 2006). These neurons and their complex interactions are the fundamental basis of all human activity – our thoughts and memories, our creative and logical abilities.

The computational complexity of a given region in the brain is dictated not only by the number of neurons it contains but also by the number of distinct cell types those neurons are divided into (Fishell and Heintz, 2013; Koch and Laurent, 1999). This raises

the obvious question of how neuron type is defined. Ideally, a combination molecular, anatomical, and physiological properties will define a given cell type (Seung and Sümbül, 2014), but variation in classification methods have led to order-of-magnitude differences in estimates of neuron types. Recent examinations of the retina, for example, have argued for as few as five neuron types (Bota et al., 2003) and as many as sixty (Masland, 2001; 2012). Such variation comes from differences in the criteria used to distinguish cell types. While some studies look only at gross morphology, spatial distribution, and connections (Bota et al., 2003), others take into account differences in gene expression, subcellular structures, and modes of intercellular signaling (Masland, 2001). The issue of defining a unique cell type is made more difficult by the relatively plastic nature of neurons. Even after a postmitotic cell acquires cell type-specific features, cellular context and activity can change the gene expression profile of the cell (Fishell and Heintz, 2013). Despite the challenge posed by this plasticity, neuron-type classification is not merely an academic exercise but instead creates a fundamental platform for asking further questions about the nervous system (Seung and Sümbül, 2014). Ultimately, a fuller understanding of neuronal cell types will come from a combination of anatomical, biochemical, and electrophysiological studies (Wichterle et al., 2013).

Because of the huge size and complexity of the vertebrate nervous system, many scientists have turned to invertebrate model organisms to understand the fundamentals of neuronal development. The nematode *Caenorhabditis elegans* has proven to be powerful system for studying the molecular mechanisms of neuron cell type determination. A small animal – only one millimeter long as an adult – *C. elegans* nevertheless exhibits a wide array of complex behaviors such as attraction to gustatory cues, mate seeking, and

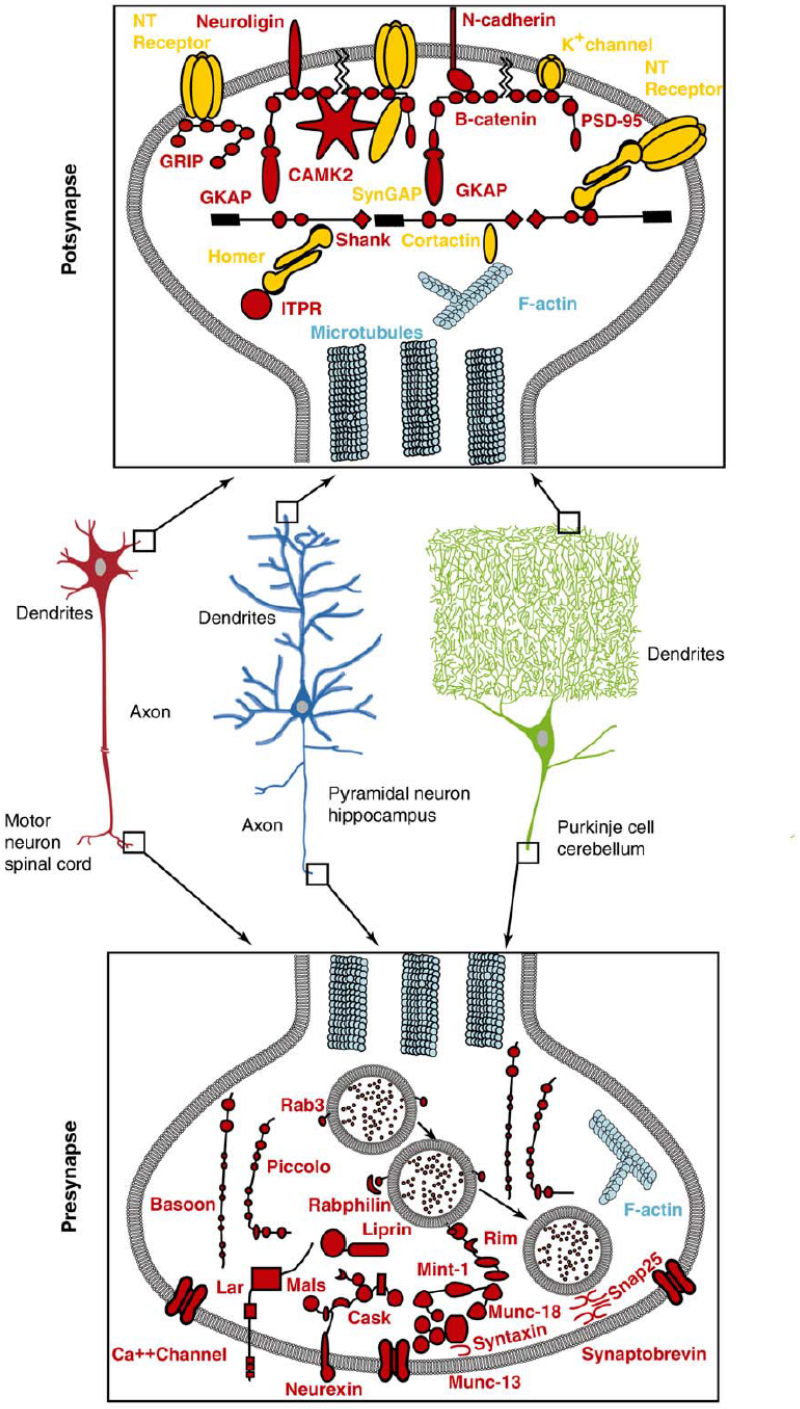
nociception response (Wood, 1988). Its fast generation time and transparent body make it easy to perform genetic screens to identify regulators of cell identity. *C. elegans* develops through an invariant cell lineage, with 957 somatic cells in adult hermaphrodites. Almost one third of those cells are neurons, which allow the animal to receive sensory information and coordinate muscle activity (Chalfie and White, 1988). The *C. elegans* nervous system has been completely mapped by electron microscopy, making it the first animal for which the entire neuronal circuitry is known (White et al., 1986). This mapping has revealed that even a simple nematode has great neuronal complexity: the neurons have traditionally been divided into 118 classes on the basis of morphology and connectivity, and recent molecular analysis has suggested that some classes should be further subdivided (Huang et al., 2004; Zhang et al., 2014). In this work, I present an analysis of the genetic programs that underlie the development of two *C. elegans* neuronal classes.

2. Pan-neuronal features

In order to attain a stable cellular identity, a neuron must undergo three steps: initial specification of neuronal features, post-mitotic establishment of cell-type identity, and refinement by local cues of characteristics such as positioning and connectivity (Fishell and Heintz, 2013). Two main cellular features are common and specific to all neuron types: cellular extensions such as axons and dendrites, and synapses, including both presynaptic and postsynaptic specializations (Figure 1) (Hobert et al., 2010). The latter includes a huge number of proteins and protein complexes. A mass spectrometry-based analysis of purified mouse postsynaptic terminals found over 1,000 proteins present in the postsynaptic proteome (Collins et al., 2006), and a *C. elegans* RNA

interference screen for actors involved in acetylcholine secretion found that 185 out of 2,072 genes screened showed secretion defects (Sieburth et al., 2005).

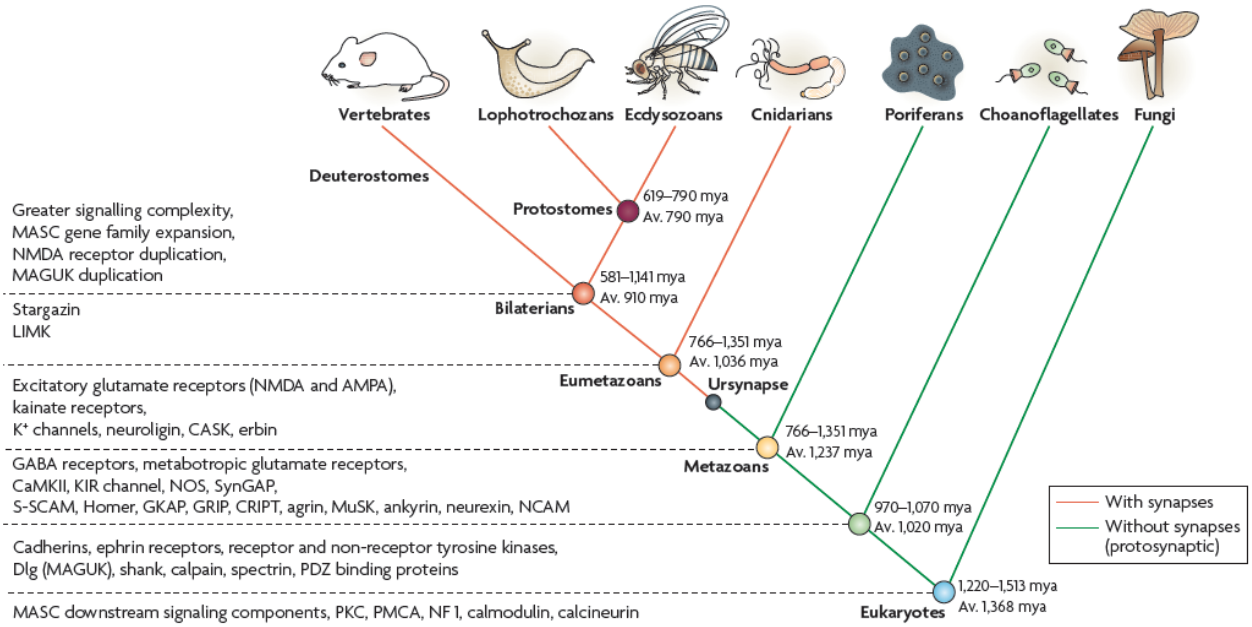
Figure 1: Common neuronal features



Three example neuronal subtypes are shown: a spinal cord motor neuron, a hippocampal pyramidal neuron, and a cerebellar Purkinje cell. While they vary in their morphology, all exhibit cellular extensions and pre- and post-synaptic specializations. Proteins shown in red are pan-neuronal, being present in most or all neuron types, while those shown in yellow are neuron-type specific (Hobert et al., 2010).

The expectation for genes that specify neuronal identity is that their expression is restricted to neurons. Surprisingly, such genes are in fact rarely expressed exclusively in neuronal cells. In many cases, presynaptic or postsynaptic proteins are found ubiquitously, with expression in every cell type tested (Hobert et al., 2010). Even simple, unicellular eukaryotes such as the yeast *Saccharomyces cerevisiae* and the amoeba *Dictyostelium discoideum* express the synaptic components PMCA and protein kinase C (Ryan and Grant, 2009). This lack of neuron type-specific regulation suggests a broader role for the synaptic machinery. Indeed, recent comparative genomics studies have suggested that proteins found in the pre-and post-synapse were already present in ancient, unicellular organisms (Figure 2). Portions of the postsynaptic density, for instance, have been found in the genomes of the unicellular choanoflagellate *Monosiga brevicollis* (Alié and Manuel, 2010) and the neuronless demosponge *Amphimedon queenslandica (Sakarya et al., 2007)*, indicating that parts of the synaptic machinery were present in the common ancestors of metazoans and these much simpler organisms. These unicellular ancestors would need to receive sensory information from the outside world, and several studies have suggested that postsynaptic complexes evolved from chemosensory modules (Achim and Arendt, 2014; Chen et al., 1999; Chiu et al., 1999).

Figure 2: Evolutionary origins of synaptic molecules



The phylogenetic tree shows the estimated time for the divergence of each clade, given in millions of years ago. The grey circle represents the ursynapse, the last common ancestor of all synapses. Various synaptic signaling molecules are shown, with the intervals at which they arose indicated. Of note, many synaptic molecules were present long before the ursynapse (Ryan and Grant, 2009).

3. Terminal fate specification

After specification of general neuronal properties, post-mitotic neurons must establish their specific neuron-type identity, a stable "ground state" that specifies cellular function (Fishell and Heintz, 2013). Heterotopic transplantation has demonstrated that identity determination requires both cell intrinsic and extrinsic cues. These experiments also suggest that specification happens through progressive restriction of neuronal fate. At rat embryonic stage E17, for instance, transplants from the neocortex to the limbic cortex fail to adopt limbic fate, while transplants between different sections of the neocortex show plasticity (Gaiano and Fishell, 1998; McConnell, 1992). Neurons, then, adopt general neocortical fate first and subsequently develop regional specificity. Similarly, when mouse E12 or P4 cerebellar progenitor cells were transplanted into the embryonic rat brain, the E12 progenitors could produce any cerebellar neuron, while the P4 progenitors were restricted to late-generated cerebellar identities (Carletti et al., 2002).

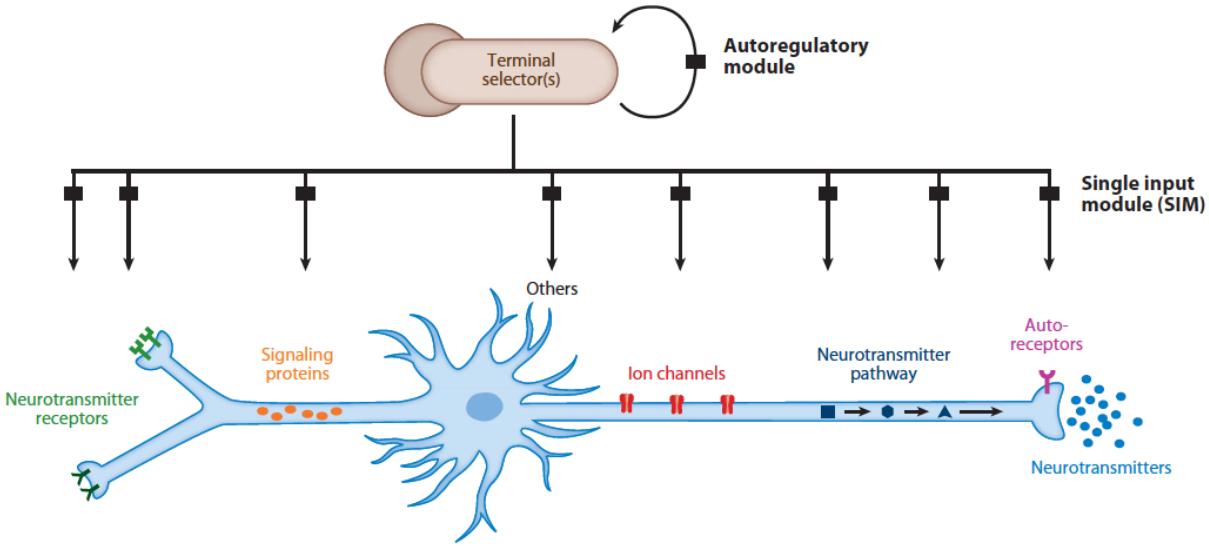
When does neuron-type specification occur? Transplantation experiments again provide insight into the timing of terminal differentiation. While exceptions exist, many neurons appear to establish their cell identity before or very shortly after becoming postmitotic (Fishell and Heintz, 2013). In an elegant experiment, ferret cortical progenitor cells were transplanted heterochronically either during the S phase of the cell cycle or later in the cell cycle. Cells transplanted from the earlier stage showed plasticity and were able to adopt the fate of their host neurons, while those cells transplanted from later in the cell cycle were already committed to their own fate (McConnell and Kaznowski, 1991). On the other hand, the homeodomain transcription factor Evx1 is expressed only postmitotically in V0 ventral spinal cord neurons, yet loss of Evx1 causes

V0 neurons to adopt a V1-like fate (Moran-Rivard et al., 2001). More work needs to be done to fully understand the temporal aspect of cell fate determination.

3.1. Terminal gene batteries

Terminally differentiated neurons can be defined by their morphology and function; these, in turn, are determined by the specific array of genes expressed in that neuron. This set of genes can collectively be referred to as the terminal gene battery (Figure 3) (Hobert, 2008; 2011; Hobert et al., 2010). These identity determinants include important neuronal routines such as ion channels, signaling proteins, synaptic adhesion molecules neurotransmitter biosynthesis pathways, receptors, and neuropeptides. The gene battery is generally considered to be stable through development and adulthood, although activity can modulate the expression of some genes (Fishell and Heintz, 2013).

Figure 3: Terminal selectors act on the terminal gene battery to define cell fate



Terminal selectors control cell fate by directly regulating the battery of genes that together define the terminally differentiated cell. In the case of a neuron, the gene battery includes molecules such as neurotransmitter receptors, signaling proteins, ion channels, neurotransmitter synthesis and packaging, cytoskeleton proteins, and adhesion molecules. Terminal selector regulate these genes through a common *cis*-regulatory motif, indicated here by the black box (Hobert, 2011).

While the complete battery of genes is unique to a particular neuron or neuron type, each individual component of the battery may be relatively broadly expressed. For example, *C. elegans* mechanosensory neurons contain unusually large microtubules composed of the alpha-tubulin MEC-12 and the beta-tubulin MEC-7. While both *mec-12* and *mec-7* are expressed in other neuron types, their expression patterns overlap only in the mechanosensory neurons (Fukushige et al., 1999).

One study to show the large differences in the terminal gene batteries of similar neurons was an analysis of two *C. elegans* sensory neuron types: the ASE gustatory neuron class and the AFD thermosensory neuron class (Etchberger et al., 2007). Using SAGE libraries of the two cell types, Etchberger and colleagues found almost six thousand genes expressed in the ASE neurons, more than one thousand of which were differentially expressed compared to AFD. These combinatorial differences can help differentiate between otherwise closely-related neuron types, as shown by a recent transcriptome analysis of the mouse retina (Siegert et al., 2012). In that study, 22 transgenic mouse lines were created with *gfp* fluorescence in different retinal cell types. Fluorescent cells were isolated from mice from each line using fluorescence-activated cell sorting, and transcripts were amplified from each. Every retina cell type was found to express a different set of genes, including different transcription factors. The transcription factors, which included genes such as *Ascl1*, *Sox2*, and *Rhox4*, are all expressed widely in vertebrates, but within the retina each was expressed in only one cell type.

3.2. The terminal selector theory

Regulation of terminal gene batteries could, in principle, be achieved in one of two ways: a) through a piecemeal activation of different components of the gene battery by many transcription factors, each responding to different cell intrinsic and extrinsic cues; or b) by cohesive activation by one transcription factor or complex that binds to a single *cis*-regulatory motif common to all genes expressed in the given cell type. Many recent studies have found support for the latter mechanism, with a single *trans*-acting factor regulating a wide array of neuron-specific genes. These factors can be referred to as terminal selectors – transcription factors which determine a given cell's fate by directly binding to and regulating its terminal gene battery (Figure 3) (Hobert, 2008; 2011).

In order to verify that a terminal selector directly binds to the regulatory regions of a neuron's selector genes, it is necessary to identify the DNA binding motif for that factor. There are several methods for accomplishing this task; in one, the *cis*-regulatory regions of genes known to be expressed in a given neuron are carefully dissected until a small sequence sufficient for expression is found. After repeating this process with many genes, the sequences are compared to identify specific motifs required for expression. This approach was used to successfully identify the binding motif for the Paired- and LIM-homeodomain proteins CEH-10 and TTX-3, which work together as terminal selectors in the *C. elegans* AIY interneuron (Wenick and Hobert, 2004). Sequential deletions were made in the promoter regions of eight AIY genes to find minimal promoter elements necessary for expression, and specific mutations were made within a 243 bp minimal promoter of the gene *ttx-3* to find disruption of AIY expression. The

mutated motif was then used to align minimal promoters of each of the other seven genes, creating a position weight matrix (PWM) that could be represented by a sequence logo (Schneider and Stephens, 1990). Similar work has been done in a variety of *C. elegans* neurons.

Given the broad expression patterns of many neuron identity genes, there must be differential regulatory motifs that allow expression in different cell types. An example of this can be seen in the *C. elegans* acytylcholine transporter *unc-17*/VAChT, an essential member of the cholinergic gene battery (Alfonso et al., 1993). *unc-17* is expressed in all cholinergic neurons, but is regulated by different transcription factors. In the DA, DB, VA, VB, and AS classes of motor neurons, *unc-17* is regulated by the COE-type transcription factor UNC-3 at the COE binding site 4187 bp upstream of the *unc-17* ATG (Kratsios et al., 2011). In the AIY interneurons, on the other hand, *unc-17* is regulated by the heterodimer of CEH-10 and TTX-3 at a binding site 2358 bp upstream of the ATG (Wenick and Hobert, 2004). These independent *cis*-regulatory motifs act in modular fashion to allow flexible regulation of any given gene in any cell type. Modularity has important evolutionary implications: not only is it comparatively easy to change the spatial and temporal regulation of a gene, but the ability to do so in only specific cells and developmental stages decreases mutational pleiotropies and increases the chances of a beneficial mutation (Wagner and Lynch, 2008).

Although terminal selectors have broad-acting effects within a cell type, they are not and cannot be the sole regulators of cell identity. Frequently, distinct cellular subroutines will be enacted by other factors acting in one of two ways: a) independently from the terminal selector; or b) as a downstream effector of the terminal selector. The

former is exemplified by the chemosensory ASE neurons, a bilateral pair of ciliated neurons that are important for chemotaxis (Bargmann and Horvitz, 1991). ASE neurons are regulated by the the zinc-finger transcription factor CHE-1, and loss of *che-1* causes a loss of ASE identity markers such as the guanylate cyclase genes *gcy-5*, *gcy-6*, and *gcy-7,* the cyclic nucleotide-gated channel *tax-2* (Uchida et al., 2003), and the glutamate transporter *eat-4* (Serrano-Saiz et al., 2013). However, the ASE neurons are still formed in *che-1* mutants, and their ciliated endings are intact (Lewis and Hodgkin, 1977). In addition, panneuronal makers such as the Ras GTPase *rab-3* are still present in ASE in *che-1* mutants (Serrano-Saiz et al., 2013). These latter findings demonstrate that ASE identity is separately regulated from pan-sensory identity and pan-neuronal identity. Indeed, the transcriptional regulation of ciliogenesis in *C. elegans* is well understood: the RFX-type transcription factor DAF-19 directly regulates many of the genes necessary for cilia biosynthesis by binding to a conserved motif known as an X box (Swoboda et al., 2000). The regulation of pan-neuronal features is not yet fully understood, but it is clear that other factors beyond the terminal selector must be acting to create the cellular extensions and synapses that characterize neurons (Hobert et al., 2010).

While terminal selectors directly regulate many if not all genes in their respective cell's gene batteries, they may also regulate additional transcription factors that act as downstream effectors of the terminal selector. Returning to the example of the ASE neurons, one of the factors regulated by the ASE terminal selector *che-1* is the homeobox gene *cog-1* (Etchberger et al., 2007). *cog-1* was identified in a search for mutants that altered the functional asymmetry between the left and right ASE neurons, and was found to act postmitotically in ASER to directly repress the LIM domain protein *lim-6*, which in

turn represses the guanylyl cyclase *gcy-5* (Chang et al., 2003). As *gcy-5* is also directly regulated by *che-1* (Etchberger et al., 2007), this represents one of the many complex interactions that occur between terminal selectors and their transcription factor targets.

3.3. Applications of the terminal selector theory in vertebrates

The terminal selector theory was first applied using examples from *C. elegans*. The abundance of cell-specific regulatory information in *C. elegans* can be attributed in part to its small number of neurons, ease of genetic manipulation and transgenesis, and fast generation time. Many studies of terminal differentiation in vertebrates have identified selector genes or master regulators whose expression is necessary for particular cell fates; the majority, however, either fail to show direct binding of the transcription factor to the regulatory sequences of the terminal gene battery or focus on regulation of just one or two genes, thus failing to show broad regulation of terminal identity markers. Nonetheless, a few excellent examples of terminal selectors have been identified in vertebrate systems, indicating the general robustness and applicability of the terminal selector theory.

One example of a terminal selector in vertebrates comes from the epidermis, a key protective layer that helps defend the body against toxic agents, dehydration, and physical stresses. In mammals, the transcription factor Kruppel-like factor 4 (Klf4) has been shown to be necessary and sufficient for the terminal differentiation of the exterior layers of the epidermis (Jaubert et al., 2003; Segre et al., 1999). To better understand the role of Klf4 on epidermal barrier formation, Patel and colleagues performed transcriptional profiling of dorsal skin before and after barrier acquisition in control, *Klf4*$^{-/-}$, and *Klf4*

overexpression mice (Patel et al., 2006). Genes that were found to be upregulated by *Klf4* included those in the epidermal differentiation complex, which code for proteins involved in the proteinaceous barrier, EGF-like ligands, and genes involved in lipid synthesis and regulation. To test for direct regulation, ~1.1 kb proximal promoter fragments of upregulated genes were cloned into luciferase constructs and transfected into a mouse epidermal cell line. When cotransfected with Klf4, 63% showed > 2-fold activation. Klf4 binding was further analyzed for the gene *Far2*. Sequential deletions were used to identify a 300 bp region containing the Klf4 binding site, and EMSA probes were used to identify smaller fragments that bound to Klf4. Finally, a specific motif was identified based on the known binding site for the erythroid Kruppel-like factor gene family (Yet et al., 1998); this motif was mutated in the luciferase assay to establish a loss of activation.

Mammalian photoreceptor cells are controlled by a variety of transcription factors, including *Crx*, *Nrl*, and *Nr2e3* (Hsiau et al., 2007). Of these, the *otx*-like homeobox transcription factor *Crx* has been shown to be expressed specifically in rod and cone cells during and after differentiation (Furukawa et al., 2002) and is necessary for proper photoreceptor formation (Furukawa et al., 1997). *Crx* binds to the *otx* family consensus binding motif, TAATCC/T. To identify the *in vivo* targets of *Crx*, ChIP-Seq was performed on 8-week-old mouse retinas (Corbo et al., 2010). Comparing the binding sites to 33 known photoreceptor *cis*-regulatory regions, it was found that 90.9% of the regions contained strong *Crx* binding sites. *Crx* binding also occurred in regions with strong phylogenetic conservation, suggesting functional importance of the binding.

Another recent example of a vertebrate terminal selector comes from the extensively characterized mouse corticospinal motor neurons (CSMN) (Jones et al., 1982). These glutamatergic subcerebral projection neurons are regulated by *Fezf2*, which is necessary for the initial specification of the cells (Molyneaux et al., 2005) as well as its axonal projections to the spinal cord and dendritic morphology (Chen et al., 2005). In an analysis by Lodato and colleagues, *Fezf2* was identified as a terminal selector for CSMNs, directing not only their glutamatergic identity but also the expression of pan-neuronal genes (Lodato et al., 2014). The authors used microarrays to search for genes upregulated in *Fezf2*-expressing cortical progenitor cells, verifying their results with *in situ* hybridization of regulated genes in control and *Fezf2*$^{-/-}$ animals. They then performed ChIP-Seq to identify Fezf2 binding sites on the proximal promoters of CSMN genes. Fezf2 was found to directly regulate the vesicular glutamate transporter *Vglut2* and axon guidance molecules such as neuropilin 2 and the tyrosine kinase receptor *Ephb1*. *Fezf2* thus represents a terminal selector gene that is able to regulate multiple neuronal subroutines.

In addition to classic terminal selectors, vertebrate systems have also produced many examples of terminal selector-like transcription factors that control specific large subroutines. For instance, vertebrate skeletal muscle cells have to respond to exercise by upregulating a wide array of genes, including those involved in angiogenesis, glucose uptake, mitochondrial biosynthesis, and reactive oxygen species detoxification. These gene programs are regulated by the peroxisome proliferator-activated receptor γ coactivator-1α, PGC-1α (Handschin, 2010). In a recent study, ChIP-Seq was performed on PGC-1α in differentiated mouse C2C12 myotubes to determine its genome-wide

binding sites (Baresic et al., 2014). The authors used the ChIP-Seq results in combination with microarray analysis performed on either control or PGC-1α overexpression cells to determine which genes were directly up- or down-regulated by PGC-1α. Significant upregulation was seen in in the vast majority of genes directly bound by PGC-1α, most of which fell into gene ontology categories related to mitochondria, oxidative phosphorylation, and energy production. Also upregulated was the transcription factor estrogen-related receptor α (ERRα), a well-known PGC-1α binding partner (Mootha et al., 2004; Schreiber et al., 2004).

4. Neurotransmitter identity

Analysis of the mechanisms of terminal differentiation and cell-type specification immediately raises the question of how cell type is defined. In their seminal analysis of the structure of the *C. elegans* nervous system, John White and colleagues used electron microscopy data to classify the 302 *C. elegans* neurons into 118 classes on the basis of gross morphology (White et al., 1986). Modern understanding of the differences in the transcriptome and proteome of individual cells, however, has provided a new way to classify cell type. One important categorization for neurons is their neurotransmitter identity, a distinction that defines the neuron's general role and also delineates its synaptic partners. Neurotransmitter type also presents a prime candidate for regulation by a terminal selector, as a core set of genes must be activated together in order for chemical signaling to occur: the neurotransmitter must be synthesized, packaged into secretory vesicles, and transported to the synapse, and following its release it must be taken back by the reuptake machinery or destroyed (Hobert et al., 2010). In the following

sections, I will give examples of terminal regulation of neurotransmitters in both *C. elegans* and vertebrates.

4.1. Dopaminergic identity

Like all members of the catecholamine family, dopamine (DA) is synthesized from tyrosine in a conserved pathway. The amino acid is hydroxylated by the enzymes CAT-2/TH and CAT-4/GCH1 and decarboxylated by BAS-1/AAAD. It is transported by the vesicular monoamine transporter CAT-1/SLC18A1/2, and its reuptake is mediated by DAT-1/SLC6A3 (Flames and Hobert, 2011). In *C. elegans,* dopamine is found in eight neurons in adult hermaphrodites, which can be dividided into four classes: ADE, CEPD, CEPV, and PDE (Chase and Koelle, 2007; Doitsidou et al., 2013). To understand the *cis*-regulatory logic of the dopamine pathway genes, Flames and Hobert dissected the proximal promoters of all five DA genes and found a small, conserved motif that was necessary for expression in the dopaminergic neurons. Using bioinformatics analysis, the regulatory motif was determined to be the binding site for an ETS transcription factor; a screen through mutants of the ten *C. elegans ets* family members identified *ast-1* as a key regulator of the dopaminergic battery, as loss of the transcription factor caused a failure in terminal differentiation of all the dopaminergic neurons. The authors then looked for ETS factors which could play a similar role in mouse dopaminergic neurons and focused on *Etv1*, which is found in all DA neurons in the olfactory bulb. Loss of *Etv1* caused a reduction in tyrosine hydroxylase-positive cells but not a reduction in neuron number of cell death, suggesting that *Etv1* is necessary for dopamine-specific differentiation. A luciferase assay in COS cells showed that Etv1 was able to activate tyrosine hydroxylase, dependent on specific dopaminergic motifs.

ast-1 and *Etv1* can not be the sole factors involved in DA neuron differentiation, however, as they are both expressed in several other neurons types besides dopaminergic neurons. Doitsidou and colleagues continued the investigation of DA-type neurons in *C. elegans* and found through further *cis*-regulatory analysis that in addition to the previously identified ETS binding site, expression in dopaminergic neurons required both a Pbx-type homeodomain binding site and a canonical homeodomain binding site (Doitsidou et al., 2013). The homeodomain transcription factor CEH-43/Dlx was identified through analysis of mutants found in a screen for loss of dopaminergic neurons. CEH-43 is necessary and in some cases sufficient to form dopaminergic neurons. To identify the Pbx-type factor, a candidate gene approach was used. Out of the three known Pbx genes in the *C. elegans* genome, *ceh-20* was found to be necessary for DA neuron differentiation in the PDE neuron class, while *ceh-20* and *ceh-40* were found to work redundantly in the other dopaminergic neuron classes.

4.2. Serotonergic identity

Like dopamine, serotonin (5-HT) is a monoamine synthesized from an aromatic amino acid, and its biosynthetic pathway shares several genes with that of dopamine (Flames and Hobert, 2011). Both neuron types require the decarboxylase BAS-1/DDC and the vesicular mononoamine transporter CAT-1/SLC18A1/2. But while the first step in dopamine synthesis is hydroxylation by the tyrosine hydroxylase CAT-2/TH, the first step in serotonin synthesis is hydroxylation by the tryptophan hydroxylase TPH-1/TPH2. It is perhaps unsurprising, then, that both neuron types are regulated by ETS domain transcription factors. In the rat brain, the *ets* gene *Pet-1* is expressed exclusively in the hindbrain serotonergic neurons, where it regulates tryptophan hydroxylase and amino

acid decarboxylase genes by binding to the conserved ETS domain binding site (Hendricks et al., 1999). *Pet-1* is also found in mice, and animals lacking *Pet-1* show deficiencies in 5-HT neuron differentiation (Hendricks et al., 2003). More recently, *Pet-1* was shown to regulate a wide range of neuronal functions in serotonergic neurons, including innervation of the somatosensory cortex and autoreceptor signaling (Liu et al., 2010).

4.3. Glutamateric identity

Unlike most neurotransmitters, glutamate is produced in every cell and so does not require any special biosynthetic enzymes. The only gene necessary to make a neuron glutamatergic, then, is a vesicular glutamate transporter, VGLUT (Takamori et al., 2000; 2001). Indeed, ectopic expression of *VGLUT* in GABAergic neurons causes synaptic release of glutamate (Takamori et al., 2000). The best-characterized VGLUT-encoding gene in *C. elegans* is *eat-4* (Lee et al., 1999), which is expressed in 78 glutamatergic neurons in adult hermaphrodites (Serrano-Saiz et al., 2013). Serrano-Saiz and colleagues dissected the proximal promoter of *eat-4* and discovered a modular regulatory logic: a series of nonoverlapping pieces of the promoter each drove expression in a specific subset of the glutamatergic neurons. Unlike dopaminergic neurons, which were all found to be controlled by the same transcription factors working together in multiple neuron classes (Doitsidou et al., 2013; Flames and Hobert, 2009), glutamatergic neuron fate is controlled by broader-acting terminal selectors (Serrano-Saiz et al., 2013). Thus *eat-4* is regulated in the ASE neurons by the ASE terminal selector *che-1*, it is regulated in ASG neurons by the ASG terminal selector *lin-11*, and so on. Mutation of specific binding

motifs in the *eat-4* locus eliminated expression in subsets of neurons that aligned with expression patterns of the corresponding transcription factor.

4.4. Cholinergic identity

Cholinergic neurons can be distinguished by the presence of four acetylcholine (ACh) pathways proteins: the choline acetyltransferase CHA-1/ChAT, which synthesizes acetylcholine, the vesicular transporter UNC-17/VAChT, the acetylcholinesterase ACE-2/AChE, which breaks down ACh into choline after synaptic release, and the choline transporter CHO-1/ChT, which allows reuptake of choline (Kratsios et al., 2011). *C. elegans* has 56 cholinergic ventral nerve cord motor neurons, which can be divided into six classes: DA-, VA-, DB-, VB-, VC-, and AS-type neurons. In all neuron types except the AS neurons, the COE-type transcription factor UNC-3 is required for expression of a wide array of genes, including the cholinergic gene battery (Kratsios et al., 2011). Loss of expression of *unc-3* causes the cells to lose their neuron-specific fate, while ectopic expression of *unc-3* causes ectopic expression of ACh pathway genes in some other neuron types. The binding site for COE-type transcription factors is known, so Kratsios and colleagues examined the promoters of genes expressed in cholinergic neurons and found that the COE motif was necessary for expression in all of the *unc-3*-dependent ventral nerve cord neurons. In addition, UNC-3 was found to directly regulate a variety of downstream effectors that repressed subtype-specific genes in certain neuron classes. For instance, the acetylcholine receptor *acr-5* is expressed only in DB- and VB-type neurons, in part because *unc-3*-mediated expression of the homeodomain transcription factor *unc-4* represses *acr-5* in the DA- and VA-type neurons.

Comparing the terminal selectors between various neurotransmitter types reveals an unexpected diversity in mechanisms. Glutamatergic fate, defined only by the presence of the vesicular glutamate transporter EAT-4/VGLUT, is regulated in a completely modular fashion in *C. elegans*. Rather than there being a small number of glutamatergic fate-inducing factors, many terminal selectors – each of which separately turns on a wide array of identity-defining genes – bind to distinct regions of the *eat-4* promoter to turn on glutamatergic fate in different cell types. Cholinergic fate, on the other hand, is governed by a limited number of terminal selector genes that each work in many cells. *unc-3/COE*, for instance, controls cholinergic fate in at least 45 neurons. In a third regulatory mechanism, dopaminergic identity requires multiple genes for complete and specific expression of the DA gene battery. These distinct means of regulating neurotransmitter type show that no basic assumptions can be made with regards to cell type specification, and more work needs to be done to understand how any given cellular subroutine is established.

5. Neuropeptidergic identity

While classic small molecule transmitters play an important role in neuronal signaling, they are not the only chemical signals that neurons use for communication. Neuropeptides have been identified throughout the animal kingdom, as well as in plants (Hökfelt et al., 2000). They have been found to play a large role in neuronal signaling pathways, mediating a wide variety of behaviors. Stress-responsive peptides are being investigated as playing a role in disorders such as anxiety and depression (Alldredge, 2010; Holmes et al., 2003; Rotzinger et al., 2010) as well as social behavior and interactions (Lukas et al., 2011; Oldfield and Hofmann, 2011). Neuropeptides exhibit

large degrees of complexity, with many families varying in their processing and chemical modulations (Hökfelt, 1991). Identification of neuropeptides has proven to be an ongoing challenge. Methods such as mass spectrometry, homology searches, and codon scanning are being used to find new peptide genes (Hummon et al., 2006; Vanden Broeck, 2001). The best-characterized neuropeptides are those of *Drosophila* and *C. elegans*, where dozens and hundreds of peptides, respectively, have been found (Li and Kyuhyung, 2008; Vanden Broeck, 2001).

One of the key distinctions between neuropeptides and classic, fast-acting neurotransmitters is the large size and high molecular weight of peptides (Salio et al., 2006). Among the implications of this difference is a much higher receptor binding affinity and specificity for neuropeptides, allowing them to signal at much lower concentrations than neurotransmitters. Additionally, peptides have long half-lives, allowing them to diffuse to non-adjacent neurons and perform long-range signaling. Indeed, evidence in frog sympathetic ganglia suggests that the peptide luteinizing-hormone-releasing hormone diffuses for tens of micrometers (Zupanc, 1996). By contrast, neurotransmitters generally have a half-life of about 5 ms, thus restricting them to signaling only at short distances (Ludwig and Leng, 2006). Neuropeptides thus represent a slower but more durable signaling option than classic neurotransmitters (Hökfelt, 1991).

5.1. Neuropeptide synthesis

Peptides are synthesized from large precursor molecules by enzymatic cleavage and chemical modification (Hook et al., 2008). Propeptides must first be cleaved by

proprotein convertases, which have varying specificities but generally cut after basic residues (Li and Kyuhyung, 2008; Rockwell et al., 2002). These cleavage sites are not well characterized in vertebrates, although computational algorithms can predict some peptides (Southey et al., 2006). One contribution to the complexity is the wide range of cleavage options: propeptides can be divided into one or more distinct neuropeptides, producing one or more copies of each (Li and Kyuhyung, 2008). Frequently, variant peptides produced by one prohormone will be identical in their C-terminal structure but will vary in their N-terminal region. These variant peptides likely have similar signaling specificities but are susceptible to different peptidases, thus increasing the probability of successful long-rang signaling (Isaac et al., 2009).

In *C. elegans*, four proprotein convertases of the Kex2/Subtilisin-like family have been identified: *kpc-1*, *egl-3*, *aex-5*, and *bli-4* (Thacker and Rose, 2000). The functional significance of one proprotein convertase, *egl-3*, was examined by analyzing the *C. elegans* peptidome using HPLC and MALDI-TOF (Husson et al., 2006). While this technique identified 75 peptides in wild type animals, a deletion mutant for *egl-3* caused a complete loss in peptide ion peaks. *egl-3* is expressed broadly in the nervous system, and mutations in the gene cause defects in mechanosensory response (Kass et al., 2001).

Following cleavage of the proprotein, the C-terminal basic residue is cleaved from the resulting peptide by carboxypeptidase E (Li and Kyuhyung, 2008). In mice, *fat* mutants – so called because they develop adult-onset obesity and hyperglycemia – have a mutation in carboxypeptidase E (Naggert et al., 1995). Loss of the enzyme causes a reduction in the processing of numerous prohormones and propeptides, including proglucagon (Friis-Hansen et al., 2001), progastrin (Udupi et al., 1997), and pro-

cholecystokini (Cain et al., 1997). The *C. elegans* gene *egl-21* encodes an orthologue of carboxypeptidase E that has been demonstrated to process FMRFamide-related protein (FLP) and neuropeptide-like protein (NLP) precursors (Husson et al., 2007; Jacob and Kaplan, 2003). Like *eg-3*, *egl-21* is broadly expressed in the nervous system (Jacob and Kaplan, 2003).

The final step in neuropeptide biosynthesis is chemical modification to inhibit degradation of the peptide signal (Li and Kyuhyung, 2008). Most commonly, the C-terminus is amidated the enzymes peptidylglycine-α-hydroxylating monooxygenase (PHM) and peptidyl-α-hydroxyglycine α-amidating lyase (PAL) (Han et al., 2004). These modifications protect the peptide from further degradation by carboxypeptidases (Isaac et al., 2009). While some candidate amidating genes have been proposed in *C. elegans*, including the monooxygenase *pamn-1*, it is not yet clear how *C. elegans* peptides are processed (Li and Kyuhyung, 2008).

5.2. Neuropeptide transport and release

Unlike classic neurotransmitters, which are usually transported in small clear vesicles, neuropeptides are processed and transported through the neuron in large dense core vesicles (LDCVs), secretory vesicles that can be identified in electron microscopy by their characteristic size and density (Ludwig and Leng, 2006; Zupanc, 1996). Once enclosed in the LDCV, peptides can be released not only at synaptic terminals like neurotransmitters, but in fact anywhere in the neuron. For example, in magnocellular neurons of the supraoptic nucleus, electron microscopy shows only a couple LDCVs in the synapse but large concentrations in the dendrites (Ludwig and Leng, 2006). This

trafficking is closely regulated. In *Aplysia*, for instance, the egg-laying hormone (ELH) prohormone is cleaved in the trans-Golgi network, and the resulting intermediate peptides are packaged into separately-sized dense core vesicles (Sossin et al., 1990). The vesicles are then transported to separate nerve terminals for release, showing that peptides from a single precursor can be differentially controlled. On the other hand, many separate peptides can also be carried together in the same LDCV, a concept known as co-storage (Merighi, 2002). Interestingly, dendritic and axonal release of LDCVs appear to be regulated separately (Ludwig and Leng, 2006). Release of both peptides and classic neurotransmitters from the synapse is controlled by Ca^{2+} entry into the axon terminus through voltage-gated ion channels. High concentrations of Ca^{2+} allow exocytosis of the small synaptic vesicles carrying neurotransmitters, which use sensors with low affinity for Ca^{2+}. Dense core vesicles, on the other hand, have a high affinity for Ca^{2+} and so can be triggered to release even upon much lower concentrations. In addition, release of neuropeptides from non-synaptic areas is not always dependent upon voltage-gated ion channels. Instead, intracellular stores of Ca^{2+} can be released, allowing exocytosis of LDCVs. In some cases, as with oxytocin and vasopressin, the neuropeptide can trigger its own release, creating a positive feedback loop that allows long-lasting signaling.

One consequence of non-synaptic release of neuropeptides is the ability of peptides to act on multiple neuronal targets. In a release known as volume transmission, neuropeptides spread diffusely over a large area of synaptic neuropil (Nässel, 2009). The peptide can then activate GPCRs in a variety of neurons, with target specificity being created not by spatial restrictions but by selectively expressed peptidases that terminate pepide action.

While only classic neurotransmitters are released from small synaptic vesicles, both small neurotransmitters and neuropeptides can be released from large dense core vesicles (Salio et al., 2006). The coexistence of neurotransmitters and neuropeptides in the same neuron has long been established (Hökfelt, 1991), but no simple rule has been found that explains which peptides neurotransmitters will interact in a given cell (Zupanc, 1996). In general, it is thought that the classic neurotransmitter is the principal messenger, with neuropeptides acting as modulators or enhancers of the signal (Salio et al., 2006). Thus the small neurotransmitter opens a ligand-gated ion channel, and the neuropeptide alters its gating properties or response to further signals, either by directly interacting with the channel or as a downstream effect of peptide-mediated activation of second messenger systems (Merighi et al., 2011). It is not clear how many, if any, neurons contain neuropeptides but not classic neurotransmitters. It has been suggested that oxytocin/vasopressin magnocellular neurons are an example of solely peptidergic neurons (Salio et al., 2006), but more recent data shows they express the type-2 vesicular glutamate transporter and thus may also be glutamatergic (Merighi et al., 2011).

5.3. *C. elegans* neuropeptide families

Neuropeptides constitute an incredibly diverse class of signaling molecules, but they can be divided into general families on the basis of sequence similarity. These include important groups such as the oxytocin/vasopressin family, the growth hormone-releasing factor superfamily, and the neuropeptide Y family (Hoyle, 1998). Three neuropeptide families have been identified in *C. elegans*: the insulin-like peptides (INS),

the FMRFamide (Phe-Met-Arg-Phe-NH_2)-related peptides (FLP), and the neuropeptide-like proteins (NLP) (Li and Kyuhyung, 2008).

Members of the insulin peptide family are secreted hormones that play diverse roles, including glucose uptake, energy storage, cell survival, and proliferation (Pierce et al., 2001). Preproinsulins consist of four domains (Pre, B, C, and A), the first two of which are cleaved off during peptide processing. The remaining domains are separated into distinct peptides which are then joined with disulfide bonds. All members of the insulin family form a similar characteristic basic loop. In *C. elegans*, an initial bioinformatics analysis of the newly-completed genome identified 37 predicted *ins*-family genes (Pierce et al., 2001); since then, an additional five genes have been identified, making the *ins* family the largest group of neuropeptides in the nematode (Li and Kim, 2010). Members of the *ins* family play a key role in development, mediating the decision to continue reproductive growth or enter dauer larval arrest (Pierce et al., 2001).

FMRFamide (Phe-Met-Arg-Phe-NH2-like) neuropeptides (FLPs) are a large and diverse family characterized by their C-terminal RFamid sequence (Li et al., 1999; McVeigh et al., 2006). At least 30 *flp* genes are present in *C. elegans*, although discrepancies between biochemical and bioinformatics analyses suggest that more may yet be found (Li et al., 1999). A study examining the expression patterns of 19 of those genes showed that *flp* peptides have distinct but overlapping expression patterns. This included expression in over 150 neurons, including all neuronal cell types. Such wide expression suggests that FLPs contribute to a variety of behaviors; however, the overlapping nature of their expression patterns indicates they may have some redundancy

in their function (Li and Kim, 2010). *flp* genes appear to be regulated by terminal selectors: three genes usually expressed in the touch receptor neurons, *flp-4*, *flp-8*, and *flp-20*, depended on the touch receptor terminal selectors MEC-3 and UNC-86 for expression (Kim and Li, 2004a).

While the *flp* and *nlp* families have well-conserved and characterized sequence and structural similarities, the final class of *C. elegans* neuropeptides is more of a "catch-all" and encompasses a diverse range of peptides. Indeed, identification of *nlp* genes was initially performed by looking for putative homologues of all known invertebrate neuropeptides (Kim and Li, 2004b). Within the large *nlp* class, several families have been identified on the basis of conserved motifs. Consistent with this variety, *nlp* genes regulate a wide array of behaviors, including fat storage, lifespan, and bacterial response (Li and Kim, 2010).

5.4. Neuropeptide regulation

While much is known about the biosynthesis and transport of neuropeptides in *C. elegans*, the regulation of neuropeptide expression is not as well understood. Neuropeptide expression is wide-spread, and most neurons have been shown to express at least one neuropeptide gene (Li and Kyuhyung, 2008). Similarly, the neuropeptide processing genes are broadly expressed, with *egl-3* being found in most neurons and *egl-21* seen in at least 60% of neurons. Neuropeptide regulation, then, lies at the level of the propeptide, but very little work has been done to understand how neuropeptide genes are limited to specific neurons in *C. elegans*.

Among the few cells in which neuropeptide regulation has been examined are the six *Drosophila* Tv neuroendocrine neurons, which are found in the embryonic ventral

nerve cord (Benveniste et al., 1998). The Tv neurons are among the approximately 60 neurons in which the *Drosophila FMRFamide* neuropeptide is expressed (Schneider et al., 1991). Expression of *FMRFamide* is modular, with different enhancers within the promoter dictating regulation in different neuron types (Benveniste et al., 1998; Schneider et al., 1993). Work by two groups has identified the regulatory cascade that controls *FMRFamide* expression in the Tv neurons (Allan et al., 2003; Benveniste et al., 1998). The LIM domain protein *apterous* (*ap*) and the zinc-finger transcription factor *squeeze* (*sqz*) act in a combinatorial manner to activate *FMRFamide* expression. While no genetic interaction between the two factors has been detected, FMRFamide levels are reduced in the double mutant compared to either single mutant alone (Allan et al., 2003). Misexpression of either factor fails to increase *FMRFamide* expression, but co-misexpression of both *ap* and *sqz* causes *FMRFamide* to be found in two other sets of neurons in the ventral nerve cord.

Part II: Homeotic Transformation

1. Interpretations of homeosis

Homeotic transformation – the change of one organ into another – has been studied and debated for more than two centuries. The term 'homeosis' (or 'homœosis') was invented by Bateson in his seminal 1894 work on the subject (Bateson, 1894), but he was not the first to describe the phenomenon. Within the field of teratology, or the study of abnormal physical development, variants that showed altered composition of some segment or organ had long been of interest. In an earlier book on vegetable teratology, Masters describes such changes as 'metamorphy', in part to avoid the misapprehensions

that had arisen by Goethe's use of the term 'metamorphosis' in 1790. "It does not necessarily imply that there has been a change in in any particular organ, but rather that there has been, to some extent, a change in the plan of construction . . . the term metamorphy is employed to distinguish cases where the ordinary course of development has been perverted or changed" (Masters, 1869). Bateson, in turn, preferred the more generally applicable 'homeosis'. While Goethe and Masters had focused on variants of plant development, such as a flower that has extra petals in the place of a stamen, Bateson emphasized the change in features within meristic, or segmented, series.

Various interpretations of homeotic transformation have historically been used, with definitions ranging from very narrow to very broad (Sattler, 1988). In the narrow sense, homeosis only applies to members of a meristic series. In the broader sense, it can refer to any change of one part of an organism to another. Sattler provides four reasons for choosing a broader interpretation over a narrower: 1) As a more general concept, the term can be applied to deal with a larger range of related phenomena. 2) As certain homologies are unclear or controversial, it can be difficult to ascertain if a transformation is homeotic or not under a narrow definition. 3) Homology itself has varying definitions, and it is unclear which should be applied. 4) Structures can be homologous on certain levels but not on others, again making interpretation of homeosis unclear. While a greater understanding of homology arising from molecular biology and genetics may alleviate some of the concerns Sattler raises, his first point remains a crucial one. Instances of developmental changes that result in one organ or tissue adopting the features of another are of great interest to evolutionary and developmental biologists, regardless of the segmental nature, or lack thereof, of the changes. Restricting the use of

the term homeosis to only a subset of transformations also inhibits comparisons of the form and mechanism of such changes. Indeed, even Bateson presents several examples of homeosis that do not meet his narrow definition of transformation within a meristic series. In a section on bilateral asymmetries, he discusses homeotic transformation between the left and right side of the body. Hens, for instance, usually have only one ovary, located on the left side. In some mutants, however, hens with two symmetrical ovaries have been found – a transformation of the right side fate to the left. Similarly, narwhal tusks usually develop from the left front tooth, but some specimens have been found with the tusk instead developing from the right tooth, or with a tusk forming from each (Bateson, 1894).

In addition to whole tissues or organs, the homeosis concept has been applied to the level of single cells. As I will discuss in later sections, lineage mutants in *C. elegans* that cause single-cell identity transformations such as *lin-12/Notch*, and *lin-22/Hairy* have historically been described as homeotic (Sternberg and Horvitz, 1984). In *Drosophila* eyes, mutations in *sevenless* that cause a transformation of the photoreceptor cell R7 into a cone cell have also been called homeotic (Tomlinson and Ready, 1986). While the homeosis terminology has not been used extensively in vertebrates, removal or ectopic expression of specific regulatory factors in many tissue types has been shown to result in cell identity switches. The transformations are often referred to as binary fate switches, but in principle, they can also be considered homeotic. In the following sections, I will review classic homeotic mutants in *Drosophila* and single cell homeosis in *C. elegans*, and I will discuss vertebrate examples of binary cell fate switches and repression of alternate identity.

2. Homeotic mutants in *Drosophila*

Classically, many examples of homeotic transformation have come from work in *Drosophila*. The segmented nature of the *Drosophila* body plan means that its homeotic mutants generally fit in the narrow definition of homeosis. *Drosophila* homeotic genes are clustered into two main regions on the right arm of the third chromosome: the Antennapedia (ANT-C) complex, which controls the formation of parasegments 1-5, and the Bithorax (BX-C) complex, which establishes the identities of the more posterior parasegments 6-14 (Noro et al., 2011). Each of these complexes is composed of several genes with related but distinct functions (Gehring and Hiromi, 1986). ANT-C is the more proximal of the two, and consists of both homeotic genes and segmental enumeration genes (Kaufman et al., 1990). The eponymous *Antennapedia* (Antp) locus promotes leg identity by repressing antennal-determining genes (Casares and Mann, 1998). As a result, null mutations in *Antp* cause the second leg to transform into an antenna (Struhl, 1981), while dominant gain-of-function mutations produce leg structures in place of antennae (Kaufman et al., 1980; Schneuwly et al., 1987b). *Antp* is normally expressed only in the posterior meso- and metathoracic segments, so its gain-of-function effects in the antenna are the result of ectopic expression (Schneuwly et al., 1987b; 1987a). Other genes in the Antennapedia complex, which includes *labial*(*lab*), *proboscipedia*(*pb*), *Deformed*(*Dfd*), and *Sex combs reduced*(*Scr*), behave in similar manners. Unusually, the location of these genes along the chromosome appears to be linked to their expression patterns in the fly (Lemons and McGinnis, 2006). Thus *lab*, the most 5' gene in the ANT-C cluster, is expressed the most anteriorly, and *Antp*, the most 3' gene, is expressed

the most posteriorly. This spatial colinearity is somewhat conserved in other organisms, although an increasing number of genomes are found to break the pattern.

The bithorax complex is located on the distal end of the right arm of chromosome 3 and consists of three genes: *Ultrabithorax*(*Ubx*), *Abdominal-A*(*AbdA*), and *Abdominal-B*(*AbdB*) (Gehring and Hiromi, 1986). Consistent with the collinear expression pattern of *Drosophila* Hox genes, members of the BX-C cluster are found in the posterior of the fly and are required for specification of the posterior thorax and abdomen. *Ubx*, the most anterior of the three, functions in parasegments 5 and 6, and mutations cause a transformation of haltere to wing – a switch to the more anterior parasegment 4 (Duncan, 1987). Gain-of-function mutations of *Ubx* cause the opposite phenotype, a transformation of the wing disc to haltere. The other two members of the BX-C complex are *AbdA* and *AbdB*, which are expressed in the abdominal region. Loss of these two genes causes a transformation of the abdomen to more anterior segments.

The predominant mechanism by which the *Drosophila* Hox genes function is through repression of alternate segment identities. Many genes required for wing formation, for example, are normally expressed in the wing disc but not the haltere. In a *Ubx* mutant, however, these genes are ectopically found in the anterior compartment of the wing disc (Weatherbee et al., 1998). To determine if the wing-determining genes were directly repressed by *Ubx*, DNase footprinting experiments were performed on *spalt*(*sal*), a DPP target known to be regulated by *Ubx*. Seven sites on the *sal* promoter were found to be protected by *Ubx*, each one containing at least one TAAT site (Galent et al., 2002). Similarly, *Antp* has been shown to repress the antenna-promoting gene *homothorax*(*hth*) (Casares and Mann, 1998).

Misexpression experiments with Hox genes from ANT-C and BX-C have revealed a surprising phenomenon: homeobox genes expressed in more posterior segments are dominant over those expressed in anterior segments (Noro et al., 2011). Thus, for example, ectopic expression of *Ultrabithorax* by a heat-shock promoter caused an anterior antenna-to-leg transformation but failed to transform posterior abdominal segments (Mann and Hogness, 1990). These consistent interactions are of interest in part because they are exposed only upon misexpression of one or more Hox genes. Under endogenous conditions, most Hox genes are expressed in distinct segments in the fly, and so a particular pattern of dominance would seem to be unnecessary. The mechanism behind posterior dominance was examined using *fkh250*, a minimal element from the *forkhead* gene that is regulated by the Antennapedia complex gene *Sex combs reduced* (*Scr*) (Noro et al., 2011). While *Fkh250* can only be bound by *Scr*, a mutated $fkh250^{CON}$ can be equally regulated by multiple Hox genes. Regulation of each promoter was tested in flies using broad coexpression of *Scr* and *AbdA*, a posterior Hox gene that acts as a transcriptional repressor. AbdA was able to win over Scr with regards to regulation of $fkh250^{CON}$ but not *fkh250*. The success of AbdA was explained by *in vitro* binding assays, which showed that AbdA had a higher binding affinity than Scr for $fkh250^{CON}$. This affinity was reliant on AbdA interactions with the Hox cofactor Extradenticle. Posterior dominance, then, can be explained at least in part by cofactor-dependent competition between anterior and posterior Hox proteins for binding to the same target genes.

3. Cell-specific fate switches in *C. elegans*

With its 957 somatic cells mapped to an invariant lineage, *C. elegans* is an ideal system for studying transformation on a single-cell level. Cell-specific fate switches are caused by changes to the cell lineage that do not alter the total number of cells but instead change the gene expression, morphology, and function of one particular cell to that of another. While *C. elegans* is not divided into strict segments like *Drosophila*, these single-cell switches have long been referred to as homeotic by analogy to the changes seen upon mutation of the *Drosophila* Hox genes (Sternberg and Horvitz, 1984).

One example of single-cell fate switches comes from the ectodermal seam cells, a row of postembryonic ectoblasts known as V1-V6 (Wrischnik and Kenyon, 1997). In wild type animals, the V1-V4 cells develop identical lineages, producing over several rounds of stem cell division a combination of epidermal cells that fuse with the epidermal syncytium and cuticle cells that form the long alae ridges. V5 and V6 cell lineages vary from the anterior four, with both producing sensory rays in the male and V5 producing a neuroblast from which a sensory structure called the postdeirid is derived. In mutants of *lin-22*, an ortholog of the *Drosophila* transcriptional regulator *hairy*, V1-V4 cells are transformed into V5 cells, with subsequent loss of epidermal nuclei and extra postdeirids and ray cells (Sternberg and Horvitz, 1984). V5-specific features are regulated by the *Antennapedia* homolog *mab-5* (Kenyon, 1986), and *mab-5* expression was found to be expanded into V1-V4 in the *lin-22* mutant (Wrischnik and Kenyon, 1997). *lin-22* thus acts as an anterior repressor of posterior fate determinants.

Within the developing *C. elegans* hermaphrodite vulva, an equivalence group is formed between two cells from neighboring sublineages: either Z1.ppp or Z4.aaa becomes an anchor cell (AC), while the other becomes a ventral uterine precursor cell

(VU). These fates undergo reciprocal homeotic transformation upon either loss or gain of expression of the Notch transmembrane receptor gene *lin-12* (Greenwald et al., 1983). In semi-dominant *lin-12*(*d*) mutants, which have elevated levels of *lin-12*, both cells become VU cells, while in *lin-12* null mutants, both cells become ACs. A similar fate switch is seen in the male, where Z1.ppp and Z4.aaa variably become a linker cell and a vas deferens precursor cell. In *lin-12*(*d*) mutants, two vas deferens cells and no linker cells are formed, while in *lin-12* null mutants, two linker cells and no vas deferens cells are formed. The effect of *lin-12* activity levels was examined using gene dosage experiments, comparing *lin-12*(*d*)/*lin-12*(+), *lin-12*(*d*)/*lin-12*(*0*), and *lin-12*(*d*)/*lin-12*(+)/*lin-12*(+). The mutant phenotype, as scored by egg-laying defects, got worse as levels of *lin-12* increased. Thus formation of the AC and VU cells is dosage-dependent, and reducing or increasing levels of *lin-12* causes reciprocal homeotic transformation of the two cell types.

Olfaction in *C. elegans* is mediated in part by three unrelated neuron pairs: AWB, which allows avoidance of repulsive odors, and AWA and AWC, which allow chemotaxis toward attractive cues (Sagasti et al., 1999). As part of a screen to find regulators of AWC neurons, Sagasti and colleagues found that mutations in the LIM homeobox gene *lim-4* caused ectopic expression of the AWC marker *str-2::gfp* in AWB neurons. Examination of the morphology of AWB in *lim-4* mutants showed that AWB cilia changed from a simple two-pronged structure to a fan-like structure similar to that normally found in AWC. Axon projections were altered as well, with many AWB axons forming an S-shape characteristic of AWC neurons. Behavioral analysis showed that in addition to gene expression and morphology, the function of AWB was also switched

such that the neuron could mediate attractive response rather than repulsive response. Upon ectopic expression of *lim-4* in the AWC neurons, the reciprocal transformation was seen: *str-2::gfp* was eliminated from AWC, while the AWB marker *str-1::gfp* was ectopically expressed in AWC.

One of the many benefits of using a powerful system like *C. elegans* to study cell fate determination is the ability to distinguish between very closely related cell types. For example, the four RME motor neurons are all GABAergic neurons that innervate head muscles to mediate foraging behavior (Huang et al., 2004). Despite their similarities and a four-fold symmetry of their cell bodies with respect to the nerve ring, they derive from different cell lineages and express different gene batteries. As such, they can be divided into two cohesive subgroups, with the lineally related RMEL and RMER comprising one group and RMED and RMEV comprising the other. These subgroups have slightly different morphology, as RMED and RMEV extend processes along the dorsal and ventral nerve cords, respectively, but RMEL and RMER do not. Work by Huang and colleagues has demonstrated that a mutation in the aryl hydrocarbon receptor *ahr-1* causes a defect in RMEL and RMER without affecting the other two RMEs. This mutant also causes changes in the morphology of RMEL and REMR, with each cell extending an ectopic process along the lateral side of the body similar to the processes normally found in RMED and RMEV. Examination of several RME subgroup-specific markers showed that the left and right RMEs were no longer expressing RMEL- and RMER-specific genes but instead were turning on genes usually only found in RMED and RMEV. *ahr-1* is usually expressed only in RMEL and RMER,

but forced expression in RMED and RMEV caused those cells to acquire RMEL/RMER-like characteristics.

A final example of cell-specific fate switches is a transformation between dorsal D (DD) and ventral D (VD) motor neurons (Shan et al., 2005; Walthall and Plunkett, 1995). These two GABAergic neuron types are formed by the action of the homeodomain protein UNC-30 (McIntire et al., 1993), and they are differentiated with the help of the nuclear receptor UNC-55, which is expressed only in the VDs (Walthall and Plunkett, 1995). In *unc-55* mutants, changes in presynaptic GABAergic varicosities indicated that the VD pre- and post-synaptic processes had switched places, making the VDs appear more DD-like. To understand the interaction between *unc-30* and *unc-55*, Shan and colleagues ectopically expressed *unc-55* in the DD neurons and found that the synaptic pattern of the resulting DDs resembled that of wild type VDs (Shan et al., 2005). They identified a DD specific reporter, *flp-13::gfp*, and found both that its expression depended on *unc-30* and that its expression was derepressed in VDs in *unc-55* mutants. They then found binding sites for both UNC-30 and UNC-55 on the *flp-13* promoter and showed that mutating the UNC-30 binding site eliminated expression of *flp-13::gfp* in the D motor neurons, while mutating the UNC-55 binding site created ectopic expression of *flp-13::gfp* in the DDs. From these results, they concluded that UNC-55 acts as a switch between DDs and VDs by repressing UNC-30-mediated activation of DD-specific genes in the VDs.

4. Repression of alternate cell fates

As an embryo undergoes development, its cells become progressively restricted in their ability to differentiate. Thus from the totipotent zygote, a series of cell divisions

produce pluripotent, then multipotent progenitor cells that finally produce individual terminally differentiated cells. This increasingly restricted state is beautifully illustrated by Waddington's epigenetic landscape, in which a cell is represented by a ball that rolls down a series of slopes, each branch further limiting the possibilities for the final destination of the ball (Goldberg et al., 2007). Mechanisms such as DNA methylation, nucleosome remodeling, and histone variants work together to make progenitor cells competent to only produce certain cell types. A terminal selector, then, can only function in particular cellular contexts. This context-dependency was shown in a study by Tursun and colleages, in which *che-1*, the terminal selector gene for the *C. elegans* ASE neurons, was ectopically expressed under a ubiquitous promoter (Tursun et al., 2011). Despite the misexpression of *che-1*, only a small subset of cells turned on the ASE gene battery. An RNAi screen of genes involved in chromatin regulation, however, showed that loss of the chromatin factor *lin-53* allowed *che-1* to turn on ASE genes in the germline. This experiment suggests that homeotic transformation upon loss or ectopic expression of a transcription factor can only occur within a given subset of cells whose epigenetic state allows them to transform.

Despite the limitations imposed by the cellular context of progenitor cells, numerous studies have found that changes in transcription factor expression result in a change in cell fate. This indicates that alternate fates are being actively repressed, either through inhibition of terminal selector genes or by direct repression of terminal gene batteries. Binary cell fate choices appear to be a recurring theme in the patterning of the nervous system, although the mechanistic basis of those choices is often not clear. In the

following sections, I will discuss the mechanisms by which cell fate is determined and examine the perturbations that lead to transformations in those cell fate decisions.

4.1. External signaling

An interesting subset of the more general phenomenon of binary cell fate switches and homeotic transformation comes from examining asymmetric divisions of progenitor cells. Many times in development, a cell divides to form two daughter cells that adopt separate fates. This asymmetry can happen in three ways, although it is important to note that these are not mutually exclusive and frequently all three mechanisms will interact to determine cell fate: 1) An intrinsic asymmetry in the mother cells that induces differential fate adoption in the daughters. This can be seen in *Drosophila* neuroblasts, in which the transcription factor Prospero accumulates on the basal side of the neuroblast and is exclusively partitioned into one daughter cell (Hirata et al., 1995). 2) lateral signaling between the daughter cells that stochastically inhibits one from adopting a particular fate, as happens in proneural clusters of *Drosophila* sensory bristles (Simpson, 1990). 3) External signaling that induces a particular terminal differentiation gene battery in one daughter cell but not the other. This last mechanism of asymmetric division has been highly studied, and a common feature of these divisions is that alterations to the signaling result in a loss of asymmetry and a adoption by one daughter cell of the fate of another.

The transmembrane receptor Notch is one of the best-characterized mediators of asymmetric divisions, and its use as a means of singling out individual cells from an equivalence group has been noted in many multicellular organisms (Jan and Jan, 1995). As noted above, mutations in *lin-12/Notch* cause reciprocal homeotic fate switches between the anchor cell and the ventral uterine precursor in *C. elegans*. In many

asymmetric divisions, interactions between Notch and the cell fate determinant Numb mediate the acquisition of asymmetry. In the *Drosophila* peripheral nervous system, for instance, external sense organs are formed from sensory organ precursors (SOPs) (Guo et al., 1996). Each SOP divides to form the daughter cells IIa and IIb, which in turn divide such that IIa produces a hair cell and a socket cell, and IIb generates a neuron and a sheath cell. In a *Notch* temperature sensitive mutant, shifting the flies to the restrictive temperature just prior to the SOP division causes the production of four neurons and none of the support cells (Hartenstein and Posakony, 1990). *numb* acts in the opposite manner as *Notch*: loss of *numb* causes the IIb cell to transform into the IIa cell, while ectopic expression of *numb* by a heat shock promoter causes the same phenotype as loss of *Notch* (Guo et al., 1996). Numb acts as a repressor of Notch signaling, preventing Notch from inducing particular fates. The interaction of Numb and Notch signaling to produce asymmetric daughter cells is dependent on the segregation of Numb into IIb, which occurs during mitosis of the progenitor cell and is dependent on *inscuteable* (Buescher et al., 1998; Jan and Jan, 1995). Notch-mediated asymmetric cell division thus requires both intrinsic and extrinsic factors.

In addition to Notch signaling, another pathway that has been shown to be involved in asymmetric cell divisions is the Wnt/β-catenin pathway. This mechanism is prevalent in *C. elegans*, where it dictates anterior-posterior cell divisions (Mizumoto and Sawa, 2007) and has been demonstrated to be active in both early embryonic divisions and terminal neuron differentiation (Bertrand and Hobert, 2009). In this pathway, stabilization of β-catenin by Wnt signal transduction allows activation of the TCF transcription factor POP-1 (Logan and Nusse, 2004). Wnt signaling is present in the

mother cell before mitosis, suggesting that Wnt induces polarity prior to cell division (Goldstein et al., 2006; Mizumoto and Sawa, 2007). In *C. elegans*, this asymmetry begins with the receptors LIN-17/Frizzled and DSH-2/Dishevelled, which are localized to the posterior side of the mother due to signaling from LIN-44/Wnt. Wnt activity in the posterior restricts other factors, such as PRY-1/Axin and APR-1/APC, to the anterior half of the cell. These factors in turn regulate the asymmetric localization of WRM-1/β-catenin. Although WRM-1 and LIT-1 are localized in the anterior of the mother cell, their position is altered during mitosis such that they accumulate mostly in the posterior daughter. There, they phosphorylate POP-1, causing it to be exported from the cell to accumulate in the nucleus of the anterior daughter. This asymmetric localization of POP-1 is then translated into changes in transcription through interaction between POP-1 and another β-catenin, SYS-1. SYS-1 is expressed in higher levels in the posterior cell, leading to the opposite expression pattern as POP-1. In the posterior cell, then, most POP-1 binds to SYS-1, converting POP-1 to an active state that allows it to induce transcription. In the posterior cell, most POP-1 is unbound and so remains in a transcriptionally repressive state. Alterations to Wnt signaling can lead to the formation of two anterior or two posterior neurons. The cholinergic interneuron AIY, for instance, differentiates from its anterior sister cell, the motor neuron SMDD, by this mechanism (Bertrand and Hobert, 2009). When a temperature sensitive Wnt-signaling mutant is moved to the restrictive temperature prior to division of the SMDD/AIY mother, the posterior daughter is transformed to the fate of the anterior and two SMDD cells are born.

4.2. Cross-antagonism between two transcription factors

A frequent theme in cell fate determination is a network of transcription factors that work antagonistically to cross-inhibit two or more cell types. In these cases, loss of either transcription factor or their regulatory network causes a de-repression of one cell type and a transformation in cell fate. A compelling example of this interaction comes from the chick dorsal horn. This area is comprised of intermingled neurons that use either GABA or glutamate as their neurotransmitter (Cheng et al., 2005). GABAergic differentiation is controlled by the LIM homeobox protein *Lhx1*, and loss of the transcription factor causes putative GABAergic neurons to instead adopt a glutamatergic fate. Glutamatergic neurons, on the other hand, are controlled by the homeobox gene *Tlx3* (Cheng et al., 2004). Loss of the *Tlx3* genes causes a switch from glutamatergic to GABAergic fate, while ectopic expression of *Tlx3* causes the opposite phenotype. Surprisingly, *Lhx1* is expressed in both GABAergic and glutamatergic neurons, and *Lhx1*$^{-/-}$ mice show no expansion of *Tlx3* expression despite the switch of neurotransmitter fate (Cheng et al., 2005). These results suggest that *Lhx1* defines a basal GABAergic state, which is somehow antagonized by *Tlx3* to induce glutamatergic fate. The means by which *Tlx3* opposes *Lhx1* function is unclear.

Molecular mechanisms have been identified in some cases of antagonistic interactions between two opposing transcription factors. In some cases, transcription factors expressed in two cells each repress expression of genes that define the opposite cell type's fate. In others, two factors expressed in the same cell compete to determine cell type, either by changes in the concentration of each factor or by physically interacting with each other to prevent functional binding. I will discuss here several

examples of known antagonistic interactions and the mechanism by which each factor represses the other.

4.2.1. Repression through physical interaction

Many of the mechanisms for antagonistic cross-inhibition are displayed in the formation of hematopoietic cells. Hematopoietic stem cells undergo differentiation into a wide range of blood cells, governed by complex interactions between the *ETS* family transcription factors PU.1, the zinc finger transcription factor GATA-1, their cofactors, and downstream effectors (Iwasaki et al., 2005). Both PU.1 and GATA-1 are expressed in early hematopoietic progenitor cells, where they compete to induce either erythroid or myeloid cell fate: enforced expression of PU.1 blocks erythroid cell differentiation, while high expression of GATA-1 prevents myeloid cells from forming (Zhang et al., 1999). Neither factor influences the expression levels of the other, suggesting that repression is due to physical interaction between the two proteins. Binding between PU.1 and GATA-1 has been demonstrated in multiple contexts. In chicken myeloblasts, cotransfection with mutants of GATA-1 showed that DNA binding by GATA-1 was not required to inhibit PU.1 action (Nerlov et al., 2000). Instead, GATA-1 was found to bind directly to PU.1 at its ETS domain. Replacing the PU.1 ETS domain with a GAL4 DNA-binding domain eliminated the ability of GATA-1 to repress myeloid cell differentiation, showing that its repressive ability require this physical interaction. The mechanism by which GATA-1-binding interferes with PU.1 function was further explained with the finding that GATA-1 competes with the PU.1 coactivator c-Jun (Zhang et al., 1999).

While repression of PU.1 by GATA-1 is relatively straightforward, studies have produced conflicting results with regards to the mechanism by which PU.1 represses GATA-1 during myelopoiesis. In one study, Zhang and colleagues found that the N-terminal transactivation domain of PU.1, but not its C-terminal DNA binding domain, was necessary for inhibition of GATA-1, and that repression of GATA-1 happened by blocking DNA binding (Zhang et al., 2000). Around the same time, Rekhtman and colleagues found that the C-terminal domain but not the N-terminal domain of PU.1 mediated interaction with the GATA-1 zinc finger domain, and that both N- and C-terminal domains were required for repression (Rekhtman et al., 1999). Several years later these seeming contradictions were resolved and the mechanism of PU.1-mediated inhibition was shown to be more complex. Binding between PU.1 and GATA-1 is dependent only on the C-terminus of PU.1 (Liew et al., 2006; Rekhtman et al., 2003), but binding alone is insufficient to inhibit GATA-1 transcriptional activity. This insufficiency suggested that some cofactor was required, and Rekhtman and colleagues found using cotransfection experiments that N-terminal binding by PU.1 of the transcriptional corepressor pRB was necessary for repression of GATA-1 (Rekhtman et al., 1999). Subsequent work found that PU.1 also recruits the histone methyltransferase Suv39h, and that pRB and Suv39h create a repressive chromatin structure with histone H3 methylation (Stopka et al., 2005). Thus cross-repression by GATA-1 and PU.1 work by two separate mechanisms: GATA-1 inhibits PU.1 by interfering with its DNA binding, while PU.1 inhibits GATA-1 by recruiting repressive complexes to it on its target genes.

4.2.2. Repression through cross-inhibition of transcription factors

In the mouse dorsal spinal cord, sensory interneurons are generated from the developing spinal neural tube. These progenitor populations show discrete, nonoverlapping expression of several bHLH factors, including Ngn1, Ngn2, and Math1 (Gowan et al., 2001). The neurons that arise from such progenitor cells can be distinguished on the basis of the transcription factors they express: those neurons that are formed from Ngn1- and Ngn2-expressing progenitors show coexpression of Lim1/2 and Brn3a, while neurons that develop from Math1-expressing progenitors are marked by LH2A/B. Single or double mutants of *Ngn1*/*Ngn2* and *Math1* do not affect the overall size of the dorsal interneuron population, but they do inhibit the terminal differentiation of each factor's respective neuron types. In particular, *Math1* mutants show a loss of LH2A/B-expressing cells and a concomitant increase in Lim1/2-expressing neurons; this change in expression was mirrored in an increase in the number of Ngn1-expressing progenitors. Conversely, *Ngn1;Ngn2* double mutants showed an expansion of the Math1 domain, although the ectopic Math1 expression was insufficient to cause an increase in LH2A/B neurons. Ectopic expression assays confirmed the repressive activity of both Ngn1 and Math1.

Another strong example of cross-repressive transcription factors comes from the mouse ventral spinal cord. There, the ventricular zone is divided into five distinct progenitor domains called p0, p1, p2, pMN, and p3 (Lee and Pfaff, 2001). These domains are distinguished by expressing different combinations of transcription factors, whose patterning is in turn established by a gradient of Shh signaling. Ten progenitor factors have been identified, and they can be divided into two classes: Class I, which are

expressed in progenitor cells unless they are repressed by Shh, and Class II, whose expression requires Shh signaling. Varying sensitivities to Shh create unique combinations of transcription factors in each of the five progenitor domains. While a morphogen gradient alone would be expected to create fuzzy boundaries between the expression domains of the TFs, in fact sharp borders exist between opposing pairs of Class I and Class II factors. Misexpression of individual transcription factors into neighboring domains shows that the proteins work cell-autonomously to inhibit activation of corresponding opposite-class factors (Briscoe et al., 2000), and this repression was found to be mediated by recruitment of Gro/TLE corepressors (Muhr et al., 2001).

4.3. Transcription factors acting as both activator and repressor

While alternative cell fates can be repressed by factors other than the terminal selector, a transcription factor in principle can simultaneously operate as an activator for some targets and a repressor for others. Indeed, many cases have been produced in which one factor appears to have this duel role. In such situations, removal of the transcription factor results in failure to activate gene batteries that define one cellular state and a derepression of gene batteries that define an alternate state. While the functional results of such manipulations are clear, the mechanism by which one factor operates in such opposing manners is often unknown. As such, it is not always possible to state definitively that activation and repression are both directly mediated by one transcription factor. Instead, one or the other function may be a downstream effect or the result of recruitment of another factor. Nonetheless, these hypothetical situations can still be distinguished from those examples I discussed in the previous section by the absence of

competing antagonistic factors. In this section, I will discuss several examples of transcription factors that act, either directly or indirectly, as both activators and repressors.

One of the most common fate switches upon loss of ectopic expression of a single transcription factor is a change in neurotransmitter identity. These switches are seen in many neurons throughout the central and peripheral nervous system, and they involve a wide variety of transcription factors. The mouse subcortical striatum, for example, is composed of distinct subpopulations of GABAergic and cholinergic interneurons (Lopes et al., 2012). Of these, the cholinergic neurons have been shown to depend on the LIM homeobox gene *Lhx7* for their differentiation (Fragkouli et al., 2005). In *Lhx7*$^{-/-}$ mice, putative cholinergic neurons express molecular markers and acquire the morphological features of GABAergic neurons (Lopes et al., 2012). This switch occurs even when *Lhx7* is removed postmitotically, and the switch to GABAergic fate does not require cell division.

A similar change in neurotransmitter fate is seen in the developing mouse mesencephalon, where neurons develop GABAergic fate over glutamatergic fate due to the action of the bHLH-O family member *Helt* (Nakatani et al., 2007). The anterior mesencephalon is divided into seven domains, m1-m7, of which *Helt* is expressed in all but the two ventral-most. Expression of *Helt* coincides with formation of GABAergic neurons, and the GABA marker *Gad1* depends on *Helt* for expression. Loss of *Helt* causes neurons in the GABA domains to instead adopt a glutamatergic fate. Notably, while *Helt* controls neurotransmitter fate, it does not affect other aspects of mesencephalon fate determination. This can particularly be seen in the neurons formed

from the m3 to m5 progeny domains. While all commit to GABAergic fate under wild type conditions, progeny from each domain vary in their expression of other transcription factors such as Nkx2.2 and Hnf3β. Although Nkx2.2 is normally expressed only in GABAergic neurons, its expression is unchanged when those neurons switch to a glutamatergic neurotransmitter type in the *Helt* mutant. The *Helt*-induced transformations thus represent only a partial fate switch and not a complete homeotic transformation, showing that particular subroutines can be activated and repressed independently of other cell identity markers.

While fate changes commonly involve switches in neurotransmitter phenotype, some factors are required to prevent other alternative fates. In the zebrafish embryonic subpallium, for example, cortical and striatal interneurons are both generated from the medial ganglionic eminence (MGE) (McKinsey et al., 2013). Both interneuron types require the *Nkx2-1* homeobox transcription factor for initial development, but *Nkx2-1* expression is quickly downregulated in cortical neurons and remains only in striatal neurons. Cortical interneurons instead require the zinc finger homeobox gene *Zfhxb1* for proper formation (Miquelajauregui et al., 2007). McKinsey and colleagues found that in the absence of *Zfhxb1*, putative cortical interneurons failed to migrate into the cortex, and the cortical markers *Cxcr7*, *Cux2*, *MafB,* and *cMaf* were strongly reduced (McKinsey et al., 2013). Concurrently, expression of *Nkx2-1* failed to be downregulated, and the cortical neurons appeared to be transformed into striatal interneurons. The mechanism of *Zfhxb1*-mediated activation and repression is unclear, although work in other neuronal contexts has demonstrated that it functions in part by recruiting repressor complexes (van Grunsven et al., 2003; Verstappen et al., 2008).

Further study of striatal neuron regulation was performed in the mouse, where the LIM homeodomain protein Islet-1(Isl1) was shown to be important in regulating differentiation of striatal neurons into striatonigral or striatopallidal cells (Lu et al., 2013). Because *Isl1* is transiently expressed, cell lineage tracing was used to determine the fate of *Isl1*-expressing cells. Immunostaining with striatonigral and striatopallidal markers showed that $Isl1^{+}$ cells developed into striatonigral neurons. In *Isl1* mutant mice, striatonigral genes were significantly downregulated and striatopallidal genes were concurrently upregulated. Similarly, ectopic expression of *Isl1* in striatopallidal neurons caused a suppression of the striatopallidal gene *Drd2*. From this, it can be concluded that *Isl1* promotes striatonigral fate while also inhibiting striatopallidal fate, but the mechanism by which it does so is unclear. *Isl1* thus represents a transcription factor that functionally appears to have the dual role of activation and repression, even though its direct role in either of those processes has not been established.

Another region of neuronal development where alternate fates must be repressed is the formation of the cerebral cortex. Cortical precursor cells express the LIM homeobox protein *Lhx2*, which has been shown to be necessary for patterning of the cerebral cortex: $Lhx2^{-/-}$ mice show loss of neocortex and an expansion of cortical hem and choroid plexus epithelium (CPe), which form from the adjacent telencephalic dorsal midline (Monuki et al., 2001). Examination of the patterning of Lhx2 and the hem marker Lmx1a and showed that the two markers have substantial overlap at E10.5 but that Lmx1a was eliminated from $Lhx2^{+}$ cells by E12.5 (Mangale et al., 2008). To determine whether this transformation occurred through a cell-autonomous role of *Lhx2*, Mangale and colleagues generated Lhx2 null mosaics. Null patches in the subpallium in

the cortex showed a loss of cortical markers and ectopic induction of hem markers, although more lateral null patches did not show hem marker expression Mutants of three downstream targets of Lhx2 were analyzed: *Foxg1*, *Emx1*, and *Pax6*. Of these, only *Foxg1* mutants showed a similar expansion of hem markers, suggesting that *Foxg1* may control the Lhx2-dependent hem fate suppression. If this is the case, then Lhx2 may act only as an activator, with its repressive functions mediated by downstream effectors.

In addition to changing individual cell fates, mutations in transcription factors can cause broad switches in regional identity. The zebrafish hindbrain, for instance, is divided into seven restricted compartments, rhombomeres 1-7. The Pbx homeodomain proteins *lzr/pbx4* and *pbx2* are expressed in an overlapping pattern in r2-r6 (Waskiewicz et al., 2002). Pbx function was eliminated in zebrafish embryos by using morpholinos against Pbx2 in *lzr/pbx4*$^{-/-}$ fish. In these animals, rhombomere identity from r2 to r6 was completely lost, while the hindbrain regions that normally would form those rhombomeres instead transformed to r1 identity. This repression of r1 identity appears to happen through interactions with *hox-1* paralogs, as morpholinos against *hoxb1a and hoxb1b strongly reduced r3 to r6 in the background of lzr/pbx4*$^{-/-}$ fish, but not in wild type animals.

While many studies have observed homeotic cell fate transformations upon gain or loss of function of a single terminal selector, very few have elucidated a mechanism by which a factor acts as both an activator and repressor. Many that do look for the means of this dual behavior have found that repression is a downstream effect of the transcription factor – either the transcription factor activates some gene that in turn acts as a repressor of alternate fate, or it recruits a repressive complex that inhibits activation

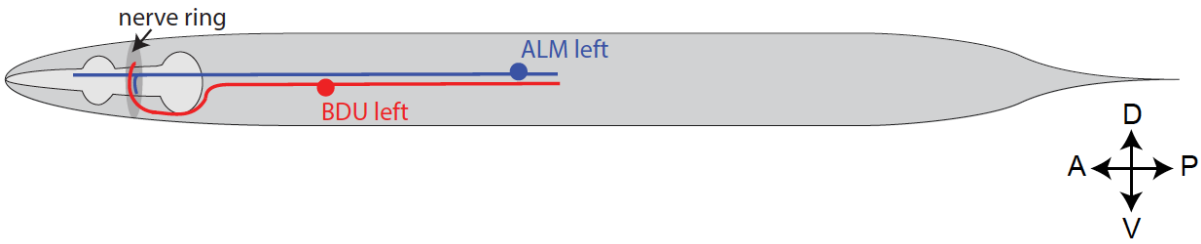
of the erroneous gene battery. In this thesis, I will present a novel mechanism by which one transcription factor, the *C. elegans* LIM-domain protein MEC-3, simultaneously acts as an activator of ALM mechanosensory neuron fate and an inhibitor of BDU interneuron fate.

Part III: The ALM/BDU Neuron Pair

1. Morphology and function

The *C. elegans* ALM and BDU neuron classes each consist of a pair of bilaterally symmetric neurons that are distinguished as 'left' or 'right'. ALM left and right (ALML/ALMR) neurons are among the six mechanosensory neurons that control the animal's response to gentle touch (Chalfie et al., 1985). The ALM cell bodies are located just in the midbody, and they extend sensory axon projections along the anterior of the animal into the head (Figure 4). The sister cells of the left and right ALM neurons, the left and right BDU neurons, are located in a more anterior lateral midbody domain of the animal. BDU extends similar anterior axons that run in a more ventral position than those of ALM; the BDU axons then turn ventrally through anterior deirid commissure just posterior of the pharynx, eventually terminating after a dorsal turn into the nerve ring (Figure 4) (White et al., 1986).

Figure 4: The ALM and BDU neurons

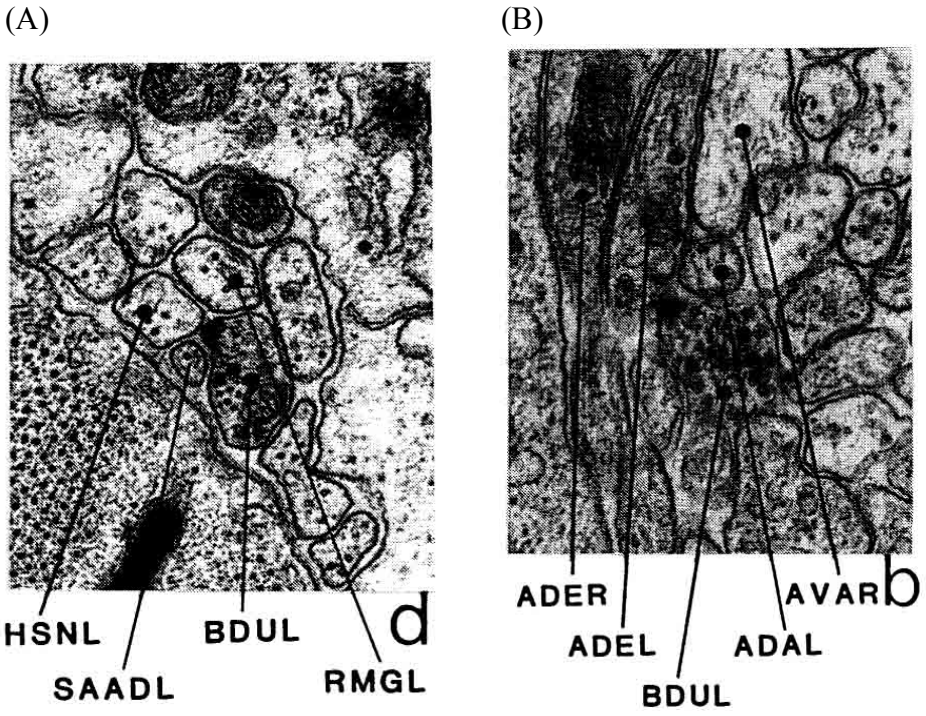


Schematic drawing of ALM and BDU. Only the left neuron of each neuron pair is shown. D, dorsal; V, ventral; A, anterior; P, posterior.

ALM and BDU neurons are generated during embryonic stages by asymmetric division of a neuroblast (Sulston et al., 1983). The sister cells are born in the anterior portion of the embryo before undergoing long-range migration that brings them to the mid-body (Hedgecock et al., 1987). The BDU cell body then migrates a short distance anteriorly as ALM continues to move posteriorly, creating a large distance between their final cell body placements (Sym et al., 1999).

Apart from their morphological differences, ALM and BDU also differ in their cellular and molecular composition. The six mechanosensory neurons are unique among *C. elegans* cells in having large microtubules with 15 protofilaments rather than the usual 11 (Fukushige et al., 1999). In addition, the two cell types differ in their choice of neurotransmitter. ALM neurons are glutamatergic (Lee et al., 1999), whereas a systematic mapping of neurotransmitter systems suggests that BDU may not use any classic, fast-acting neurotransmitter (Duerr et al., 2001; McIntire et al., 1993; Serrano-Saiz et al., 2013) (our unpublished data). Instead, in contrast to most *C. elegans* neurons, all of the synaptic endings of BDU contain striking, darkly staining vesicles (Figure 5), suggesting that the BDU neurons make prominent use of neuropeptides (White et al., 1986). Indeed, five neuropeptide-producing genes are known to be expressed in BDU: *flp-10*, *flp-12*, *nlp-1*, *nlp-15*, and *nlp-37*. None of these genes are expressed in ALM, although other neuropeptides are, including *flp-20* (Kim and Li, 2004a; Li et al., 1999).

Figure 5:



Electron microscopy images were made of BDU and other neurons. In contrast to the surrounding neurons, BDU contains darkly stained vesicles that characterize neuropeptide release. Many of these vesicles are not at the synapse (personal communication with David Hall), indicating non-synaptic release (White et al., 1986).

Functionally, ALM and BDU are quite different. ALM neurons, along with the other four mechanosensory neurons, mediate the animal's response to gentle touch (Chalfie et al., 1985). They receive sensory input from direct mechanical gating of voltage-independent DEG/ENaC Na^+ channels composed of the proteins MEC-4 and MEC-10, as well as the accessory subunits MEC-2 and MEC-6 (O'Hagan et al., 2005). BDU neurons, in contrast, are generally referred to as interneurons because they lack any highly branched or ciliated endings, or other morphological markers that could identify them as sensory neurons. However, recent cell ablation studies have demonstrated that the BDU neurons are involved in the harsh touch circuitry in the anterior half of the animals (Li et al., 2011). It is unclear whether BDU neurons directly act as mechanoreceptors or whether they act downstream of some other sensory neuron. However, the long processes that extend from BDU along the anterior half of the animal suggest that it may indeed be mechanosensory.

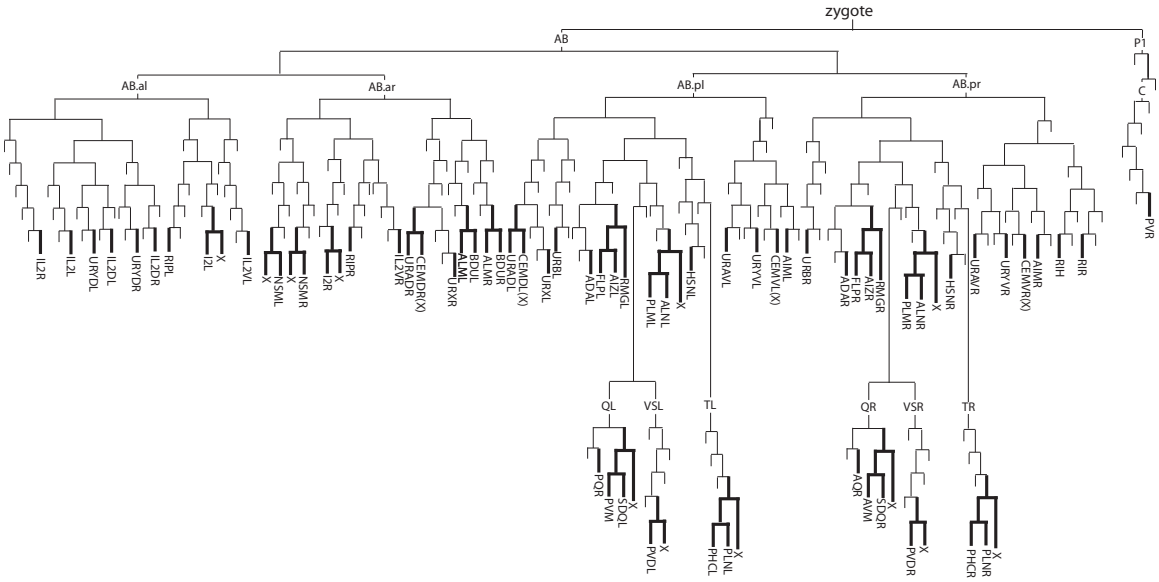
2. Regulation of ALM and BDU identity

Many years of work have gone in to studying the ALM and other mechanosensory neurons, and as a result their regulation is well understood. Two transcription factors work together as terminal selectors in ALM: The POU domain transcription factor UNC-86 and the LIM homeobox protein MEC-3 (Duggan et al., 1998).

Expression of *unc-86* in the embryo is restricted to neurons and neuroblasts (Figure 6). It is expressed in 57 neurons in adult hermaphrodites, including the ALM and BDU neurons (Baumeister et al., 1996; Finney et al., 1988). UNC-86 acts as a terminal selector in many of those cells, including the serotonergic NSM and HSN neurons and

the cholinergic IL2 sensory and URA motor neurons (Sze et al., 2002; Zhang et al., 2014). Such diversity of cell types all controlled by one transcription factor indicates that UNC-86 works in combination with cofactors that help specify these different cell fates. This is indeed the case: in addition to interacting with MEC-3 in the ALM cells, *unc-86* has been shown to interact with factors such as the LIM homeodomain protein TTX-3 and the ARID-type transcription factor CFI-1.

Figure 6: *unc-86* expression patterns



A limited *C. elegans* lineage showing expression *unc-86* expression. Dark lines indicate cells where *unc-86* is expressed. Adapted from Finney and Ruvkun, 1990.

The variable interactions between UNC-86 and different cofactors in different cell types are consistent with what is known about POU domain proteins from vertebrate and *in vitro* studies. The POU domain – named as an acronym for the vertebrate proteins Pit-1, Oct-1, and Oct-2, and the *C. elegans* protein UNC-86 – consists of a structurally bipartite domain in which a POU-specific box (POU_S) and POU homeobox (POU_{HD}) are separated by a non-conserved short linker region (Herr et al., 1988; Phillips and Luisi, 2000). The linker creates flexible binding to DNA, allowing the protein to bind to different DNA motifs in different cellular contexts (Ryan and Rosenfeld, 1997). In particular, DNA binding affinity and specificity are frequently affected by interactions between POU domain proteins and their cofactors. In a crystal structure of a ternary complex between the POU domain protein Oct1, an octomer DNA site, and the coactivator OCA-B, the cofactor was found to stabilize Oct1 on the DNA by binding to an exposed hydrophobic pocket on the POU-specific domain (Chasman et al., 1999). Similarly, dynamic NMR analysis of Oct-1 found that in a binary Oct1-DNA complex, rapid exchange occurred between multiple configurations such that either the POU_S or the POU_{HD} domains were unbound to the DNA (Williams et al., 2004). When a cofactor, Sox2, was added to the complex, a dramatic reduction in dissociation of the POU domain from the DNA was seen.

In ALM and the other mechanosensory neurons, UNC-86 activates MEC-3 by binding to several conserved sequences on the *mec-3* promoter (Way et al., 1991; Xue et al., 1992). MEC-3 then interacts with UNC-86 on its own promoter, creating a positive feedback loop for *mec-3* expression (Way and Chalfie, 1989; Xue et al., 1992). In addition, the UNC-86/MEC-3 heterodimer acts as a terminal selector for the whole

mechanosensory gene battery by directly activating genes needed to respond to gentle touch and signal to other neurons (Duggan et al., 1998; Serrano-Saiz et al., 2013; Zhang et al., 2002). Loss of either *unc-86* or *mec-3* causes a mechanosensory defect, with animals unable to respond to gentle touch (Chalfie and Au, 1989; Way and Chalfie, 1988).

In vitro studies show that while UNC-86 is able to bind on its own to the *mec-3* promoter, the addition of MEC-3 greatly increases its binding affinity due to a decrease in the dissociation rate of the heterodimer (Xue et al., 1993) and causes a synergistic increase in transcriptional output (Lichtsteiner and Tjian, 1995). The interactions between UNC-86 and MEC-3 are important for their *in vivo* function, as a mutant that dramatically decreased MEC-3 DNA binding but increased heterodimer formation was able to rescue a *mec-3* mutant (Xue et al., 1993). Heterodimer formation is mediated by the POU-specific domain of UNC-86 (Röckelein et al., 2000; Röhrig et al., 2000) and a region C-terminal of the MEC-3 homeodomain (Xue et al., 1993).

Less is known about the regulation of BDU cell fate than that of ALM. While *unc-86* is expressed in BDU as well as ALM, it is unclear whether it acts as a terminal selector in both neuron classes. One factor that has been shown to affect BDU development is the zinc-finger transcription factor PAG-3/GFI-1. First isolated in a screen for ectopic expression of *mec-7lacZ*, *pag-3* mutants were found to cause several genes normally expressed only in ALM to be expressed in BDU as well, although *mec-3* was not found to be ectopically expressed (Jia et al., 1996). Moreover, *pag-3* negatively regulates its own expression, as Northern analysis showed an increase in *pag-3* transcripts in a *pag-3* null mutant (Jia et al., 1997). Surprisingly, *pag-3* is expressed not

just in BDU but also in ALM and other mechanosensory neurons. Its role in ALM is not clear.

The repressive role that *pag-3* appears to play by inhibiting the expression of ALM genes in BDU correlates with the known function of the *Drosophila* ortholog *senseless* (*sens*). *Senseless* is expressed in the R8 photoreceptor and acts to prevent alternative identity formation at several stages of development. Loss of *sens* at early stages causes presumptive R8 cells to instead adopt R2/R5 fate (Frankfort et al., 2001). This fate transformation is due to ectopic expression of the transcriptional repressor *rough*, which normally acts in R2 and R5 to repress R8 fate. Ectopic expression of *pag-3* causes loss of endogenous *rough* expression and formation of ectopic R8 photoreceptors. During later stages of photoreceptor development, *senseless* inhibits R8 acquisition of R7 features by repressing the R7 opsins Rh3 and Rh4 through a conserved binding site (Nolo et al., 2000; Xie et al., 2007).

While *senseless* has most frequently been found to act as a transcriptional repressor, evidence suggests that it can also assume the role of a transcriptional activator, given the correct cellular and promoter context. In some cases, this has been found to be a concentration-dependent effect. Low levels of Senseless repress transcription of the bHLH factor *achaete*, while high levels activate it (Jafar-Nejad and Bellen, 2004). Similarly, the mammalian PAG-3 homolog Gfi-1 has been shown to either activate or repress genes in erythroid cell lines, depending on the promoter (Jafar-Nejad and Bellen, 2004). Given the dual role of the PAG-3 homologs, it is unclear how PAG-3 is functioning in ALM and BDU.

References

Achim, K., and Arendt, D. (2014). Structural evolution of cell types by step-wise assembly of cellular modules. Current Opinion in Genetics and Development *27*, 102–108.

Alfonso, A., Grundahl, K., Duerr, J.S., Han, H.P., and Rand, J.B. (1993). The *Caenorhabditis elegans* unc-17 gene: a putative vesicular acetylcholine transporter. Science *261*, 617–619.

Alié, A., and Manuel, M. (2010). The backbone of the post-synaptic density originated in a unicellular ancestor of choanoflagellates and metazoans. BMC Evolutionary Biology *10*, 34.

Allan, D.W., St Pierre, S.E., Miguel-Aliaga, I., and Thor, S. (2003). Specification of Neuropeptide Cell Identityby the Integration of Retrograde BMP Signaling and a Combinatorial Transcription Factor Code. Cell *113*, 73–86.

Alldredge, B. (2010). Pathogenic involvement of neuropeptides in anxiety and depression. Neuropeptides *44*, 215–224.

Andres-Barquin, P.J. (2002). Santiago Ramón y Cajal and the Spanish school of neurology. The Lancet Neurology *1*, 445–452.

Azevedo, F.A.C., Carvalho, L.R.B., Grinberg, L.T., Farfel, J.M., Ferretti, R.E.L., Leite, R.E.P., Jacob Filho, W., Lent, R., and Herculano-Houzel, S. (2009). Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. J. Comp. Neurol. *513*, 532–541.

Baresic, M., Salatino, S., Kupr, B., van Nimwegen, E., and Handschin, C. (2014). Transcriptional Network Analysis in Muscle Reveals AP-1 as a Partner of PGC-1α in the Regulation of the Hypoxic Gene Program. Molecular and Cellular Biology *34*, 2996–3012.

Bargmann, C.I., and Horvitz, H.R. (1991). Chemosensory neurons with overlapping functions direct chemotaxis to multiple chemicals in *C. elegans*. Neuron *7*, 729–742.

Bateson, W. (1894). Materials for the Study of Variation (London: Macmillan and Co).

Baumeister, R., Liu, Y., and Ruvkun, G. (1996). Lineage-specific regulators couple cell lineage asymmetry to the transcription of the *Caenorhabditis elegans* POU gene unc-86 during neurogenesis. Genes & Development *10*, 1395–1410.

Benveniste, R.J., Thor, S., Thomas, J.B., and Taghert, P.H. (1998). Cell type-specific regulation of the *Drosophila* FMRF-NH2 neuropeptide gene by Apterous, a LIM homeodomain transcription factor. Development *125*, 4757–4765.

Bertrand, V., and Hobert, O. (2009). Linking Asymmetric Cell Division to the Terminal

Differentiation Program of Postmitotic Neurons in *C. elegans*. Developmental Cell *16*, 563–575.

Bota, M., Dong, H.-W., and Swanson, L.W. (2003). From gene networks to brain networks. Nat Neurosci *6*, 795–799.

Briscoe, J., Pierani, A., Jessell, T.M., and Ericson, J. (2000). A Homeodomain Protein Code Specifies Progenitor Cell Identity and Neuronal Fate in the Ventral Neural Tube. Cell *101*, 435–445.

Buescher, M., Yeo, S.L., Udolph, G., Zavortink, M., Yang, X., Tear, G., and Chia, W. (1998). Binary sibling neuronal cell fate decisions in the Drosophila embryonic central nervous system are nonstochastic and require inscuteable-mediated asymmetry of ganglion mother cells. Genes & Development *12*, 1858–1870.

Cain, B.M., Wang, W., and Beinfeld, M.C. (1997). Cholecystokinin (CCK) Levels Are Greatly Reduced in the Brains But Not the Duodenums of Cpefat/Cpefat Mice: A Regional Difference in the Involvement of Carboxypeptidase E (Cpe) in Pro-CCK Processing. Endocrinology *138*, 4034–4037.

Carletti, B., Grimaldi, P., Magrassi, L., and Rossi, F. (2002). Specification of Cerebellar Progenitors After Heterotopic–Heterochronic Transplantation to the Embryonic CNSIn Vivo and In Vitro. Journal of Neuroscience *22*, 7132–7146.

Casares, F., and Mann, R.S. (1998). Control of antennal versus leg development in *Drosophila*. Nature *392*, 723–726.

Chalfie, M., and Au, M. (1989). Genetic control of differentiation of the *Caenorhabditis elegans* touch receptor neurons. Science *243*, 1027–1033.

Chalfie, M., and White, J.G. (1988). The Nervous System. In The Nematode *Caenorhabditis Elegans*, pp. 337–392.

Chalfie, M., Sulston, J.E., White, J.G., Southgate, E., Thomson, J.N., and Brenner, S. (1985). The Neural Circuit for Touch Sensitivity in *Caenorhabditis elegans*. Journal of Neuroscience *5*, 956–964.

Chang, S., Robert J Johnston, J., Johnston, and Hobert, O. (2003). A transcriptional regulatory cascade that controls left/right asymmetry in chemosensory neurons of *C. elegans*. Genes & Development *17*, 2123–2137.

Chase, D.L., and Koelle, M.R. (2007). Biogenic amine neurotransmitters in *C. elegans*. In WormBook, E.M. Jorgensen, and J.M. Kaplan, eds.

Chasman, D., Cepek, K., Sharp, P.A., and Pabo, C.O. (1999). Crystal structure of an OCA-B peptide bound to an Oct-1 POU domain/octamer DNA complex: specific recognition of a protein-DNA interface. Genes & Development *13*, 2650–2657.

Chen, G.-Q., Cui, C., Mayer, M.L., and Gouaux, E. (1999). Functional characterization of a potassium-selective prokaryotic glutamate receptor. Nature *402*, 817–821.

Chen, J.-G., Rašin, M.-R., Kwan, K.Y., and Šestan, N. (2005). Zfp312 is required for subcortical axonal projections and dendritic morphology of deep-layer pyramidal neurons of the cerebral cortex. Proceedings of the National Academy of Sciences *102*, 17792–17797.

Cheng, L., Arata, A., Mizuguchi, R., Qian, Y., Karunaratne, A., Gray, P.A., Arata, S., Shirasawa, S., Bouchard, M., Luo, P., et al. (2004). Tlx3 and Tlx1 are post-mitotic selector genes determining glutamatergic over GABAergic cell fates. Nat Neurosci *7*, 510–517.

Cheng, L., Samad, O.A., Xu, Y., Mizuguchi, R., Luo, P., Shirasawa, S., Goulding, M., and Ma, Q. (2005). Lbx1 and Tlx3 are opposing switches in determining GABAergic versus glutamatergic transmitter phenotypes. Nat Neurosci *8*, 1510–1515.

Chiu, J., DeSalle, R., Lam, H.-M., Meisel, L., and Coruzzi, G. (1999). Molecular evolution of glutamate receptors: a primitive signaling mechanism that existed before plants and animals diverged. Mol Biol Evol *16*, 826–838.

Collins, M.O., Husi, H., Yu, L., Brandon, J.M., Anderson, C.N.G., Blackstock, W.P., Choudhary, J.S., and Grant, S.G.N. (2006). Molecular characterization and comparison of the components and multiprotein complexes in the postsynaptic proteome. J. Neurochem. *97 Suppl 1*, 16–23.

Corbo, J.C., Lawrence, K.A., Karlstetter, M., Myers, C.A., Abdelaziz, M., Dirkes, W., Weigelt, K., Seifert, M., Benes, V., Fritsche, L.G., et al. (2010). CRX ChIP-seq reveals the cis-regulatory architecture of mouse photoreceptors. Genome Res. *20*, 1512–1525.

Doitsidou, M., Flames, N., Topalidou, I., Abe, N., Felton, T., Remesal, L., Popovitchenko, T., Mann, R.S., Chalfie, M., and Hobert, O. (2013). A combinatorial regulatory signature controls terminal differentiation of the dopaminergic nervous system in *C. elegans*. Genes & Development *27*, 1391–1405.

Duerr, J.S., Gaskin, J., and Rand, J.B. (2001). Identified neurons in *C. elegans* coexpress vesicular transporters for acetylcholine and monoamines. American Journal of Physiology - Cell Physiology *280*, C1616–C1622.

Duggan, A., Ma, C., and Chalfie, M. (1998). Regulation of touch receptor differentiation by the *Caenorhabditis elegans mec-3* and *unc-86* genes. Development *125*, 4107–4119.

Duncan, I. (1987). The Bithorax Complex. Annual Review of Genetics *21*, 285–319.

Etchberger, J.F., Lorch, A., Sleumer, M.C., Zapf, R., Jones, S.J., Marra, M.A., Holt, R.A., Moerman, D.G., and Hobert, O. (2007). The molecular signature and cis-regulatory architecture of a *C. elegans* gustatory neuron. Genes & Development *21*, 1653–1674.

Finney, M., Ruvkun, G., and Horvitz, H.R. (1988). The *C. elegans* Cell Lineage and Differentiation Gene *unc-86* Encodes a Protein with a Homeodomain and Extended Similarity to Transcription Factors. Cell *55*, 757–769.

Fishell, G., and Heintz, N. (2013). The Neuron Identity Problem: Form Meets Function. Neuron *80*, 602–612.

Flames, N., and Hobert, O. (2009). Gene regulatory logic of dopamine neuron differentiation. Nature *458*, 885–889.

Flames, N., and Hobert, O. (2011). Transcriptional Control of the Terminal Fate of Monoaminergic Neurons. Annu. Rev. Neurosci. *34*, 153–184.

Fragkouli, A., Hearn, C., Errington, M., Cooke, S., Grigoriou, M., Bliss, T., Stylianopoulou, F., and Pachnis, V. (2005). Loss of forebrain cholinergic neurons and impairment in spatial learning and memory in LHX7-deficient mice. European Journal of Neuroscience *21*, 2923–2938.

Frankfort, B.J., Nolo, R., Zhang, Z., Bellen, H., and Mardon, G. (2001). *senseless* Repression of *rough* is Required for R8 Photoreceptor Differentiation in the Developing *Drosophila* Eye. Neuron *32*, 403–414.

Friis-Hansen, L., Lacourse, K.A., Samuelson, L.C., and Holst, J.J. (2001). Attenuated processing of proglucagon and glucagon-like peptide-1 in carboxypeptidase E-deficient mice. J Endocrinol *169*, 595–602.

Fukushige, T., Siddiqui, Z.K., Chou, M., Culotti, J.G., Gogonea, C.B., Siddiqui, S.S., and Hamelin, M. (1999). MEC-12, an alpha-tubulin required for touch sensitivity in *C. elegans*. J Cell Sci *112 (Pt 3)*, 395–403.

Furukawa, A., Koike, C., Lippincott, P., Cepko, C.L., and Furukawa, T. (2002). The Mouse Crx 5′-Upstream Transgene Sequence Directs Cell-Specific and Developmentally Regulated Expression in Retinal Photoreceptor Cells. Journal of Neuroscience *22*, 1640–1647.

Furukawa, T., Morrow, E.M., and Cepko, C.L. (1997). Crx, a Novel otx-like Homeobox Gene, Shows Photoreceptor-Specific Expression and Regulates Photoreceptor Differentiation. Cell *91*, 531–541.

Gaiano, N., and Fishell, G. (1998). Transplantation as a tool to study progenitors within the vertebrate nervous system. J. Neurobiol. *36*, 152–161.

Galent, R., Walsh, C.M., and Carroll, S.B. (2002). Hox repression of a target gene: extradenticle-independent, additive action through multiple monomer binding sites. Development *129*, 3115–3126.

Gehring, W.J., and Hiromi, Y. (1986). Homeotic genes and the homeobox. Annual Review of Genetics *20*, 147–173.

Goldberg, A.D., Allis, C.D., and Bernstein, E. (2007). Epigenetics: A Landscape Takes Shape. Cell *128*, 635–638.

Goldstein, B., Takeshita, H., Mizumoto, K., and Sawa, H. (2006). Wnt Signals Can Function as Positional Cues in Establishing Cell Polarity. Developmental Cell *10*, 391–396.

Gowan, K., Helms, A.W., Hunsaker, T.L., Collisson, T., Ebert, P.J., Odom, R., and Johnson, J.E. (2001). Crossinhibitory Activities of Ngn1 and Math1 Allow Specificationof Distinct Dorsal Interneurons. Neuron *31*, 219–232.

Greenwald, I.S., Sternberg, P.W., and Horvitz, H.R. (1983). The *lin-12* locus specifies cell fates in Caenorhabditis elegans. Cell *34*, 435–444.

Guo, M., Jan, L.Y., and Jan, Y.N. (1996). Control of Daughter Cell Fates during Asymmetric Division: Interaction of Numb and Notch. Neuron *17*, 27–41.

Han, M., Park, D., Vanderzalm, P.J., Mains, R.E., Eipper, B.A., and Taghert, P.H. (2004). *Drosophila* uses two distinct neuropeptide amidating enzymes, dPAL1 and dPAL2. J. Neurochem. *90*, 129–141.

Handschin, C. (2010). Regulation of skeletal muscle cell plasticity by the peroxisome proliferator-activated receptor γ coactivator 1α. Journal of Receptors and Signal Transduction *30*, 376–384.

Hartenstein, V., and Posakony, J.W. (1990). A dual function of the Notch gene in *Drosophila* sensillum development. Developmental Biology *142*, 13–30.

Hedgecock, E.M., Culotti, J.G., Hall, D.H., and Stern, B.D. (1987). Genetics of cell and axon migrations in *Caenorhabditis elegans*. Development *100*, 365–382.

Hendricks, T.J., Fyodorov, D.V., Wegman, L.J., Lelutiu, N.B., Pehek, E.A., Yamamoto, B., Silver, J., Weeber, E.J., Sweatt, J.D., and Deneris, E.S. (2003). Pet-1 ETS Gene Plays a Critical Role in 5-HT Neuron Development and Is Required for Normal Anxiety-like and Aggressive Behavior. Neuron *37*, 233–247.

Hendricks, T., Francis, N., Fyodorov, D., and Deneris, E.S. (1999). The ETS domain factor Pet-1 is an early and precise marker of central serotonin neurons and interacts with a conserved element in serotonergic genes. Journal of Neuroscience *19*, 10348–10356.

Herr, W., Sturm, R.A., Clerc, R.G., Corcoran, L.M., Baltimore, D., Sharp, P.A., Ingraham, H.A., Rosenfeld, M.G., Finney, M., and Ruvkun, G. (1988). The POU domain: a large conserved region in the mammalian *pit-1, oct-1, oct-2*, and *Caenorhabditis elegans unc-86* gene products. Genes & Development *2*, 1513–1516.

Hirata, J., Nakagoshi, H., Nabeshima, Y.-I., and Matsuzaki, F. (1995). Asymmetric segregation of the homeodomain protein Prospero during *Drosophila* development. Nature *377*, 627–630.

Hobert, O. (2008). Regulatory logic of neuronal diversity: Terminal selector genes and selector motifs. Proceedings of the National Academy of Sciences *105*, 20067–20071.

Hobert, O. (2011). Regulation of Terminal Differentiation Programs in the Nervous System. Annu. Rev. Cell Dev. Biol. *27*, 681–696.

Hobert, O., Carrera, I., and Stefanakis, N. (2010). The molecular and gene regulatory signature of a neuron. Trends in Neurosciences *33*, 435–445.

Holmes, A., Heilig, M., Rupniak, N.M.J., Steckler, T., and Griebel, G. (2003). Neuropeptide systems as novel therapeutic targets for depression and anxiety disorders. Trends in Pharmacological Sciences *24*, 580–588.

Hook, V., Funkelstein, L., Lu, D., Bark, S., Wegrzyn, J., and Hwang, S.-R. (2008). Proteases for Processing Proneuropeptides into Peptide Neurotransmitters and Hormones. Annual Review of Pharmacology and Toxicology *48*, 393–423.

Hoyle, C.H.V. (1998). Neuropeptide families: evolutionary perspectives. Regulatory Peptides *73*, 1–33.

Hökfelt, T. (1991). Neuropeptides in Perspective: Review The Last Ten Years. Neuron *7*, 867–879.

Hökfelt, T., Broberger, C., Xu, Z.-Q.D., Sergeyev, V., Ubink, R., and Diez, M. (2000). Neuropeptides — an overview. Neuropharmacology *39*, 1337–1356.

Hsiau, T.H.C., Diaconu, C., Myers, C.A., Lee, J., Cepko, C.L., and Corbo, J.C. (2007). The *Cis*-regulatory Logic of the Mammalian Photoreceptor Transcriptional Network. PLoS ONE *2*, e643.

Huang, X., Powell-Coffman, J.A., and Jin, Y. (2004). The AHR-1 aryl hydrocarbon receptor and its co-factor the AHA-1 aryl hydrocarbon receptor nuclear translocator specify GABAergic neuron cell fate in *C. elegans*. Development *131*, 819–828.

Hummon, A.B., Richmond, T.A., Verleyen, P., Baggerman, G., Huybrechts, J., Ewing, M.A., Vierstraete, E., Rodriguez-Zas, S.L., Schoofs, L., Robinson, G.E., et al. (2006). From the genome to the proteome: uncovering peptides in the Apis brain. Science *314*, 647–649.

Husson, S.J., Clynen, E., Baggerman, G., Janssen, T., and Schoofs, L. (2006). Defective processing of neuropeptide precursors in *Caenorhabditis elegans* lacking proprotein convertase 2 (KPC-2/EGL-3): mutant analysis by mass spectrometry. J. Neurochem. *98*, 1999–2012.

Husson, S.J., Janssen, T., Baggerman, G., Bogert, B., Kahn Kirby, A.H., Ashrafi, K., and Schoofs, L. (2007). Impaired processing of FLP and NLP peptides in carboxypeptidase E (EGL-21)-deficient *Caenorhabditis elegans* as analyzed by mass spectrometry. J. Neurochem. *102*, 246–260.

Isaac, R.E., Bland, N.D., and Shirras, A.D. (2009). Neuropeptidases and the metabolic inactivation of insect neuropeptides. General and Comparative Endocrinology *162*, 8–17.

Iwasaki, H., Somoza, C., Shigematsu, H., Duprez, E.A., Iwasaki-Arai, J., Mizuno, S.-I., Arinobu, Y., Geary, K., Zhang, P., Dayaram, T., et al. (2005). Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. Blood *106*, 1590–1600.

Jacob, T.C., and Kaplan, J.M. (2003). The EGL-21 Carboxypeptidase E Facilitates Acetylcholine Release at Caenorhabditis elegans Neuromuscular Junctions. Journal of Neuroscience *23*, 2122–2130.

Jafar-Nejad, H., and Bellen, H.J. (2004). Gfi/Pag-3/Senseless Zinc Finger Proteins: a Unifying Theme? Molecular and Cellular Biology *24*, 8803–8812.

Jan, Y.N., and Jan, L.Y. (1995). Maggot"s hair and bug"s eye: Role of cell interactions and intrinsic factors in cell fate specification. Neuron *14*, 1–5.

Jaubert, J., Cheng, J., and Segre, J.A. (2003). Ectopic expression of kruppel like factor 4 (Klf4) accelerates formation of the epidermal permeability barrier. Development *130*, 2767–2777.

Jia, Y., Xie, G., McDermott, J.B., and Aamodt, E. (1997). The *C. elegans* gene pag-3 is homologous to the zinc finger proto-oncogene gfi-1. Development *124*, 2063–2073.

Jia, Y., Xie, G., and Aamodt, E. (1996). *pag-3*, a *Caenorhabditis elegans* Gene Involved in Touch Neuron Gene Expression and Coordinated Movement. Genetics *142*, 141–147.

Jones, E.G., Schreyer, D.J., and Wise, S.P. (1982). Growth and Maturation of the Rat Corticospinal Tract. (Elsevier), pp. 361–379.

Kass, J., Jacob, T.C., Kim, P., and Kaplan, J.M. (2001). The EGL-3 Proprotein Convertase Regulates Mechanosensory Responses of *Caenorhabditis elegans*. Journal of Neuroscience *21*, 9265–9272.

Kaufman, T.C., Lewis, R., and Wakimoto, B. (1980). CYTOGENETIC ANALYSIS OF CHROMOSOME 3 IN *DROSOPHILA MELANOGASTER*: THE HOMOEOTIC GENE COMPLEX IN POLYTENE CHROMOSOME INTERVAL 84A-B. Genetics *94*, 115–133.

Kaufman, T.C., Seeger, M.A., and Olsen, G. (1990). Molecular and Genetic Organization of The Antennapedia Gene Complex of *Drosophila melanogaster*. In Genetic Regulatory Hierarchies in Development, (Elsevier), pp. 309–362.

Kenyon, C. (1986). A gene involved in the development of the posterior body region of *C. elegans*. Cell *46*, 477–487.

Kim, K., and Li, C. (2004a). Expression and regulation of an FMRFamide-related

neuropeptide gene family in *Caenorhabditis elegans*. J. Comp. Neurol. *475*, 540–550.

Kim, K., and Li, C. (2004b). Expression and regulation of an FMRFamide-related neuropeptide gene family in *Caenorhabditis elegans*. J. Comp. Neurol. *475*, 540–550.

Koch, C., and Laurent, G. (1999). Complexity and the nervous system. Science *284*, 96–98.

Kratsios, P., Stolfi, A., Levine, M., and Hobert, O. (2011). Coordinated regulation of cholinergic motor neuron traits through a conserved terminal selector gene. Nat Neurosci *15*, 205–214.

Lee, R.Y.N., Sawin, E.R., Chalfie, M., Horvitz, H.R., and Avery, L. (1999). EAT-4, a Homolog of a Mammalian Sodium-Dependent Inorganic Phosphate Cotransporter, Is Necessary for Glutamatergic Neurotransmission in *Caenorhabditis elegans*. Journal of Neuroscience *19*, 159–167.

Lee, S.-K., and Pfaff, S.L. (2001). Transcriptional networks regulating neuronal identity in the developing spinal cord. Nat Neurosci *4*, 1183–1191.

Lemons, D., and McGinnis, W. (2006). Genomic Evolution of Hox Gene Clusters. Science *313*, 1918–1922.

Lewis, J.A., and Hodgkin, J.A. (1977). Specific neuroanatomical changes in chemosensory mutants of the nematode *Caenorhabditis elegans*. J. Comp. Neurol. *172*, 489–510.

Li, C., and Kim, K. (2010). Neuropeptide Gene Families in *Caenorhabditis elegans*. In Neuropeptide Systems as Targets for Parasite and Pest Control, T.G. Geary, and A.G. Maule, eds. pp. 98–137.

Li, C., and Kyuhyung, K. (2008). Neuropeptides. In Wormbook.

Li, C., Kim, K., and Nelson, L.S. (1999). FMRFamide-related neuropeptide gene family in *Caenorhabditis elegans*. Brain Research *848*, 26–34.

Li, W., Kang, L., Piggott, B.J., Feng, Z., and Xu, X.Z.S. (2011). The neural circuits and sensory channels mediating harsh touch sensation in *Caenorhabditis elegans*. Nat Comms *2*, 315.

Lichtsteiner, S., and Tjian, R. (1995). Synergistic activation of transcription by UNC-86 and MEC-3 in *Caenorhabditis elegans* embryo extracts. The EMBO Journal *14*, 3937–3945.

Liew, C.W., Rand, K.D., Simpson, R.J.Y., Yung, W.W., Mansfield, R.E., Crossley, M., Proetorius-Ibba, M., Nerlov, C., Poulsen, F.M., and Mackay, J.P. (2006). Molecular Analysis of the Interaction between the Hematopoietic Master Transcription Factors GATA-1 and PU.1. Journal of Biological Chemistry *281*, 28296–28306.

Liu, C., Maejima, T., Wyler, S.C., Casadesus, G., Herlitze, S., and Deneris, E.S. (2010). *Pet-1* is required across different stages of life to regulate serotonergic function. Nat Neurosci *13*, 1190–1198.

Lodato, S., Molyneaux, B.J., Zuccaro, E., Goff, L.A., Chen, H.-H., Yuan, W., Meleski, A., Takahashi, E., Mahony, S., Rinn, J.L., et al. (2014). Gene co-regulation by Fezf2 selects neurotransmitter identity and connectivity of corticospinal neurons. Nat Neurosci *17*, 1046–1054.

Logan, C.Y., and Nusse, R. (2004). The Wnt signaling pathway in development and disease. Annu. Rev. Cell Dev. Biol. *20*, 781–810.

Lopes, R., van Wijk, N.V., Neves, G., and Pachnis, V. (2012). Transcription factor LIM homeobox 7 (Lhx7) maintainssubtype identity of cholinergic interneurons in themammalian striatum. Proceedings of the National Academy of Sciences *109*, 3119–3124.

López-Muñoz, F., Boya, J., and Alamo, C. (2006). Neuron theory, the cornerstone of neuroscience, on the centenary of the Nobel Prize award to Santiago Ramón y Cajal. Brain Research Bulletin *70*, 391–405.

Lu, K.M., Evans, S.M., Hirano, S., and Liu, F.C. (2013). Dual role for *Islet-1* in promoting striatonigral and repressing striatopallidal genetic programs to specify striatonigral cell identity. Proceedings of the National Academy of Sciences *111*, E168–E177.

Ludwig, M., and Leng, G. (2006). Dendritic peptide release and peptide-dependent behaviours. Nat Rev Neurosci *7*, 126–136.

Lukas, M., Toth, I., Reber, S.O., Slattery, D.A., Veenema, A.H., and Neumann, I.D. (2011). The Neuropeptide Oxytocin Facilitates Pro-Social Behavior and Prevents Social Avoidance in Rats and Mice. Neuropsychopharmacology *36*, 2159–2168.

Mangale, V.S., Hirokawa, K.E., Satyaki, P.R.V., Gokulchandran, N., Chikbire, S., Subramanian, L., Shetty, A.S., Martynoga, B., Paul, J., Mai, M.V., et al. (2008). Lhx2 Selector Activity Specifies Cortical Identity and Suppresses Hippocampal Organizer Fate. Science *319*, 304–309.

Mann, R.S., and Hogness, D.S. (1990). Functional dissection of ultrabithorax proteins in D. melanogaster. Cell *60*, 597–610.

Masland, R.H. (2001). The fundamental plan of the retina. Nat Neurosci *4*, 877–886.

Masland, R.H. (2012). The Neuronal Organization of the Retina. Neuron *76*, 266–280.

Masters, M.T. (1869). Vegetable Teratology (London: Robert Hardwicke).

McConnell, S.K., and Kaznowski, C.E. (1991). Cell cycle dependence of laminar

determination in developing neocortex. Science *254*, 282–285.

McConnell, S.K. (1992). The control of neuronal identity in the developing cerebral cortex. Current Opinion in Neurobiology *2*, 23–27.

McIntire, S.L., Jorgensen, E., and Horvitz, H.R. (1993). Genes required for GABA function in *Caenorhabditis elegans*. Nature *364*, 334–337.

McKinsey, G.L., Lindtner, S., Trzcinski, B., Visel, A., Pennacchio, L.A., Huylebroeck, D., Higashi, Y., and Rubenstein, J.L.R. (2013). Dlx1&2-Dependent Expression of Zfhx1b (Sip1, Zeb2) Regulates the Fate Switch between Cortical and Striatal Interneurons. Neuron *77*, 83–98.

McVeigh, P., Geary, T.G., Marks, N.J., and Maule, A.G. (2006). The FLP-side of nematodes. Trends in Parasitology *22*, 385–396.

Merighi, A. (2002). Costorage and coexistence of neuropeptides in the mammalian CNS. Progress in Neurobiology *66*, 161–190.

Merighi, A., Salio, C., Ferrini, F., and Lossi, L. (2011). Neuromodulatory function of neuropeptides in the normal CNS. Journal of Chemical Neuroanatomy *42*, 276–287.

Miquelajauregui, A., Van de Putte, T., Polyakov, A., Nityanandam, A., Boppana, S., Seuntjens, E., Karabinos, A., Higashi, Y., Huylebroeck, D., and Tarabykin, V. (2007). Smad-interacting protein-1 (Zfhx1b) acts upstream of Wnt signaling in the mouse hippocampus and controls its formation. Proceedings of the National Academy of Sciences *104*, 12919–12924.

Mizumoto, K., and Sawa, H. (2007). Two βs or not two βs: regulation of asymmetric division by β-catenin. Trends in Cell Biology *17*, 465–473.

Molyneaux, B.J., Arlotta, P., Hirata, T., Hibi, M., and Macklis, J.D. (2005). Fezl Is Required for the Birth and Specification of Corticospinal Motor Neurons. Neuron *47*, 817–831.

Monuki, E.S., Porter, F.D., and Walsh, C.A. (2001). Patterning of the Dorsal Telencephalon and Cerebral Cortex by a Roof Plate-Lhx2 Pathway. Neuron *32*, 591–604.

Mootha, V.K., Handschin, C., Arlow, D., Xie, X., St Pierre, J., Sihag, S., Yang, W., Altshuler, D., Puigserver, P., Patterson, N., et al. (2004). Erralpha and Gabpa/b specify PGC-1alpha-dependent oxidative phosphorylation gene expression that is altered in diabetic muscle. Proceedings of the National Academy of Sciences *101*, 6570–6575.

Moran-Rivard, L., Kagawa, T., Saueressig, H., Gross, M.K., Burrill, J., and Goulding, M. (2001). Evx1 Is a Postmitotic Determinant of V0 Interneuron Identity in the Spinal Cord. Neuron *29*, 385–399.

Muhr, J., Andersson, E., Persson, M., Jessell, T.M., and Ericson, J. (2001). Groucho-

Mediated Transcriptional Repression Establishes Progenitor Cell Pattern and Neuronal Fate in the Ventral Neural Tube. Cell *104*, 861–873.

Muotri, A.R., and Gage, F.H. (2006). Generation of neuronal variability and complexity. Nature *441*, 1087–1093.

Naggert, J.K., Fricker, L.D., Varlamov, O., Nishina, P.M., Rouille, Y., Steiner, D.F., Carroll, R.J., Paigen, B.J., and Leiter, E.H. (1995). Hyperproinsulinaemia in obese fat/fat mice associated with a carboxypeptidase E mutation which reduces enzyme activity. Nature Genetics *10*, 135–142.

Nakatani, T., Minaki, Y., Kumai, M., and Ono, Y. (2007). Helt determines GABAergic over glutamatergic neuronal fate by repressing Ngn genes in the developing mesencephalon. Development *134*, 2783–2793.

Nässel, D.R. (2009). Neuropeptide signaling near and far: how localized and timed is the action of neuropeptides in brain circuits? Invert Neurosci *9*, 57–75.

Nerlov, C., Querfurth, E., Kulessa, H., and Graf, T. (2000). GATA-1 interacts with the myeloid PU.1 transcription factor and represses PU.1-dependent transcription. Blood *95*, 2543–2551.

Nolo, R., Abbott, L.A., and Bellen, H.J. (2000). Senseless, a Zn Finger Transcription Factor, Is Necessary and Sufficient for Sensory Organ Development in *Drosophila*. Cell *102*, 349–362.

Noro, B., Lelli, K., Sun, L., and Mann, R.S. (2011). Competition for cofactor-dependent DNA binding underlies Hox phenotypic suppression. Genes & Development *25*, 2327–2332.

O'Hagan, R., Chalfie, M., and Goodman, M.B. (2005). The MEC-4 DEG/ENaC channel of *Caenorhabditis elegans* touch receptor neurons transduces mechanical signals. Nat Neurosci *8*, 43–50.

Oldfield, R.G., and Hofmann, H.A. (2011). Neuropeptide regulation of social behavior in a monogamous cichlid fish. Physiology & Behavior *102*, 296–303.

Pablo Garcia-Lopez, V.G.-M.M.F. (2010). The Histological Slides and Drawings of Cajal. Frontiers in Neuroanatomy *4*, 9.

Patel, S., Xi, Z.F., Seo, E.Y., McGaughey, D., and Segre, J.A. (2006). Klf4 and corticosteroids activate an overlapping set of transcriptional targets to accelerate in utero epidermal barrier acquisition. Proceedings of the National Academy of Sciences *103*, 18668–18673.

Phillips, K., and Luisi, B. (2000). The virtuoso of versatility: POU proteins that flex to fit. Journal of Molecular Biology *302*, 1023–1039.

Piccolino, M., Strettoi, E., and Laurenzi, E. (1989). Santiago Ramon y Cajal, the retina and the neuron theory. Documenta Ophtalmologica *71*, 123–141.

Pierce, S.B., Costa, M., Wisotzkey, R., Devadhar, S., Homburger, S.A., Buchman, A.R., Ferguson, K.C., Heller, J., Platt, D.M., Pasquinelli, A.A., et al. (2001). Regulation of DAF-2 receptor signaling by human insulin and *ins-1*, a member of the unusually large and diverse *C. elegans* insulin gene family. Genes & Development *15*, 672–686.

Rekhtman, N., Choe, K.S., Matushansky, I., Murray, S., Stopka, T., and Skoultchi, A.I. (2003). PU.1 and pRB Interact and Cooperate To Repress GATA-1 and Block Erythroid Differentiation. Molecular and Cellular Biology *23*, 7460–7474.

Rekhtman, N., Radparvar, F., and Evans, T. (1999). Direct interaction of hematopoietic transcription factors PU.1 and GATA-1: functional antagonism in erythroid cells. Genes & Development *13*, 1398–1411.

Rockwell, N.C., Krysan, D.J., Komiyama, T., and Fuller, R.S. (2002). Precursor Processing by Kex2/Furin Proteases. Chemical Reviews *102*, 4525–4548.

Rotzinger, S., Lovejoy, D.A., and Tan, L.A. (2010). Behavioral effects of neuropeptides in rodent models of depression and anxiety. Peptides *31*, 736–756.

Röckelein, I., Röhrig, S., Donhauser, R., Eimer, S., and Baumeister, R. (2000). Identification of Amino Acid Residues in the *Caenorhabditis elegans* POU Protein UNC-86 That Mediate UNC-86-MEC-3-DNA Ternary Complex Formation. Molecular and Cellular Biology *20*, 4806–4813.

Röhrig, S., Röckelein, I., Donhauser, R., and Baumeister, R. (2000). Protein interaction surface of the POU transcriptionfactor UNC-86 selectively used in touch neurons. The EMBO Journal *19*, 3694–3703.

Ryan, A.K., and Rosenfeld, M.G. (1997). POU domain family values: flexibility, partnerships, and developmental codes. Genes & Development *11*, 1207–1225.

Ryan, T.J., and Grant, S.G.N. (2009). The origin and evolution of synapses. Nat Rev Neurosci *10*, 701–712.

Sagasti, A., Hobert, O., Troemel, E.R., Ruvkun, G., and Bargmann, C.I. (1999). Alternative olfactory neuron fates are specified by the LIM homeobox gene. Genes & Development *13*, 1794–1806.

Sakarya, O., Armstrong, K.A., Adamska, M., Adamski, M., Wang, I.-F., Tidor, B., Degnan, B.M., Oakley, T.H., and Kosik, K.S. (2007). A post-synaptic scaffold at the origin of the animal kingdom. PLoS ONE *2*, e506.

Salio, C., Lossi, L., Ferrini, F., and Merighi, A. (2006). Neuropeptides as synaptic transmitters. Cell Tissue Res. *326*, 583–598.

Sattler, R. (1988). Homeosis in Plants. American Journal of Botany *75*, 1606–1617.

Schneider, L.E., O'Brien, M.A., and Taghert, P.H. (1991). In situ hybridization analysis of the FMRFamide neuropeptide gene in Drosophila. I. Restricted expression in embryonic and larval stages. J. Comp. Neurol. *304*, 608–622.

Schneider, L.E., Roberts, M.S., and Taghert, P.H. (1993). Cell type-specific transcriptional regulation of the *drosophila* FMRFamide neuropeptide gene. Neuron *10*, 279–291.

Schneider, T.D., and Stephens, R.M. (1990). Sequence logos: a new way to display consensus sequences. Nucl. Acids Res. *18*, 6097–6100.

Schneuwly, S., Klemenz, R., and Gehring, W.J. (1987a). Redesigning the body plan of *Drosophila* by ectopic expression of the homoeotic gene *Antennapedia*. Nature *325*, 816–818.

Schneuwly, S., Kuroiwa, A., and Gehring, W.J. (1987b). Molecular analysis of the dominant homeotic Antennapedia phenotype. The EMBO Journal *6*, 201.

Schreiber, S.N., Emter, R., Hock, M.B., Knutti, D., Cardenas, J., Podvinec, M., Oakeley, E.J., and Kralli, A. (2004). The estrogen-related receptor alpha (ERRalpha) functions in PPARgamma coactivator 1alpha (PGC-1alpha)-induced mitochondrial biogenesis. Proceedings of the National Academy of Sciences *101*, 6472–6477.

Segre, J.A., Bauer, C., and Fuchs, E. (1999). Klf4 is a transcription factor required for establishing the barrier function of the skin. Nature Genetics *22*, 356–360.

Serrano-Saiz, E., Poole, R.J., Felton, T., Zhang, F., La Cruz, De, E.D., and Hobert, O. (2013). Modular Control of Glutamatergic Neuronal Identity in *C. elegans* by Distinct Homeodomain Proteins. Cell *155*, 659–673.

Seung, H.S., and Sümbül, U. (2014). Neuronal Cell Types and Connectivity: Lessons from the Retina. Neuron *83*, 1262–1272.

Shan, G., Kim, K., Li, C., and Walthall, W.W. (2005). Convergent genetic programs regulate similarities and differences between related motor neuron classes in *Caenorhabditis elegans*. Developmental Biology *280*, 494–503.

Sieburth, D., Ch'ng, Q., Dybbs, M., Tavazoie, M., Kennedy, S., Wang, D., Dupuy, D., Rual, J.-F., Hill, D.E., Vidal, M., et al. (2005). Systematic analysis of genes required for synapse structure and function. Nature *436*, 510–517.

Siegert, S., Cabuy, E., Scherf, B.G., Kohler, H., Panda, S., Le, Y.-Z., Fehling, H.J., Gaidatzis, D., Stadler, M.B., and Roska, B. (2012). Transcriptional code and disease map for adult retinal cell types. Nat Neurosci *15*, 487–95–S1–2.

Simpson, P. (1990). Lateral inhibition and the development of the sensory bristles of the

adult peripheral nervous system of Drosophila. Development *109*, 509–519.

Sossin, W.S., Sweet-Cordero, A., and Scheller, R.H. (1990). Dale's hypothesis revisited: different neuropeptides derived from a common prohormone are targeted to different processes. Proceedings of the National Academy of Sciences *87*, 4845–4848.

Southey, B.R., Amare, A., Zimmerman, T.A., Rodriguez-Zas, S.L., and Sweedler, J.V. (2006). NeuroPred: a tool to predict cleavage sites in neuropeptide precursors and provide the masses of the resulting peptides. Nucl. Acids Res. *34*, W267–W272.

Sternberg, P.W., and Horvitz, H.R. (1984). The Genetic Control of Cell Lineage During Nematode Development. Annual Review of Genetics *18*, 489–524.

Stopka, T., Amanatullah, D.F., Papetti, M., and Skoultchi, A.I. (2005). PU.1 inhibits the erythroid program by binding to GATA-1 on DNA and creating a repressive chromatin structure. The EMBO Journal *24*, 3712–3723.

Struhl, G. (1981). A homoeotic mutation transforming leg to antenna in *Drosophila*. Nature *292*, 635–638.

Sulston, J.E., Schierenberg, E., White, J.G., and Thomson, J.N. (1983). The Embryonic Cell Lineage of the Nematode *Caenorhabditis elegans*. Developmental Biology *100*, 64–119.

Swoboda, P., Adler, H.T., and Thomas, J.H. (2000). The RFX-Type Transcription Factor DAF-19 Regulates Sensory Neuron Cilium Formation in *C. elegans*. Molecular Cell *5*, 411–421.

Sym, M., Robinson, N., and Kenyon, C. (1999). MIG-13 Positions Migrating Cells along the Anteroposterior Body Axis of. Cell *98*, 25–36.

Sze, J.Y., Zhang, S., Li, J., and Ruvkun, G. (2002). The *C. elegans* POU-domain transcription factor UNC-86 regulates the tph-1 tryptophan hydroxylase gene and neurite outgrowth in specific serotonergic neurons. Development *129*, 3901–3911.

Takamori, S., Rhee, J.S., Rosenmund, C., and Jahn, R. (2000). Identification of a vesicular glutamate transporter that defines a glutamatergic phenotype in neurons. Nature *407*, 189–194.

Takamori, S., Rhee, J.S., Rosenmund, C., and Jahn, R. (2001). Identification of Differentiation-Associated Brain-Specific Phosphate Transporter as a Second Vesicular Glutamate Transporter (VGLUT2). Journal of Neuroscience *21*, 1–6.

Thacker, C., and Rose, A.M. (2000). A look at the *Caenorhabditis elegans* Kex2/Subtilisin-like proprotein convertase family. Bioessays *22*, 545–553.

Tomlinson, A., and Ready, D.F. (1986). Sevenless: A Cell-Specific Homeotic Mutation of the Drosophila Eye. Science *231*, 400–402.

Tursun, B., Patel, T., Kratsios, P., and Hobert, O. (2011). Direct conversion of C. elegans germ cells into specific neuron types. Science *331*, 304–308.

Uchida, O., Nakano, H., Koga, M., and Ohshima, Y. (2003). The *C. elegans* che-1 gene encodes a zinc finger transcription factor required for specification of the ASE chemosensory neurons. Development *130*, 1215–1224.

Udupi, V., Gomez, P., Song, L., Varlamov, O., Reed, J.T., Leiter, E.H., Fricker, L.D., and George H Greeley, J. (1997). Effect of Carboxypeptidase E Deficiency on Progastrin Processing and Gastrin Messenger Ribonucleic Acid Expression in Mice with the fat Mutation1. Endocrinology *138*, 1959–1963.

van Grunsven, L.A., Michiels, C., Van de Putte, T., Nelles, L., Wuytens, G., Verschueren, K., and Huylebroeck, D. (2003). Interaction between Smad-interacting protein-1 and the corepressor C-terminal binding protein is dispensable for transcriptional repression of E-cadherin. Journal of Biological Chemistry *278*, 26135–26145.

Vanden Broeck, J. (2001). Neuropeptides and their precursors in the fruitfly, *Drosophila melanogaster*. Peptides *22*, 241–254.

Verstappen, G., van Grunsven, L.A., Michiels, C., Van de Putte, T., Souopgui, J., Van Damme, J., Bellefroid, E., Vandekerckhove, J., and Huylebroeck, D. (2008). Atypical Mowat-Wilson patient confirms the importance of the novel association between ZFHX1B/SIP1 and NuRD corepressor complex. Hum. Mol. Genet. *17*, 1175–1183.

Wagner, G.P., and Lynch, V.J. (2008). The gene regulatory logic of transcription factor evolution. Trends in Ecology & Evolution *23*, 377–385.

Walthall, W.W., and Plunkett, J.A. (1995). Genetic transformation of the synaptic pattern of a motoneuron class in Caenorhabditis elegans. Journal of Neuroscience *15*, 1035–1043.

Waskiewicz, A.J., Rikhof, H.A., and Moens, C.B. (2002). Eliminating Zebrafish Pbx Proteins Reveals a Hindbrain Ground State. Developmental Cell *3*, 723–733.

Way, J.C., and Chalfie, M. (1989). The mec-3 gene of *Caenorhabditis elegans* requires its own product for maintained expression and is expressed in three neuronal cell types. Genes & Development *3*, 1823–1833.

Way, J.C., Wang, L., Run, J.Q., and Wang, A. (1991). The mec-3 gene contains cis-acting elements mediating positive and negative regulation in cells produced by asymmetric cell division in Caenorhabditis elegans. Genes & Development *5*, 2199–2211.

Way, J.C., and Chalfie, M. (1988). *mec-3*, a homeobox-containing gene that specifies differentiation of the touch receptor neurons in *C. elegans*. Cell *54*, 5–16.

Weatherbee, S.D., Halder, G., Kim, J., Hudson, A., and Carroll, S. (1998). Ultrabithorax

regulates genes at several levels of the wing-patterning hierarchy to shape the development of the *Drosophila* haltere. Genes & Development *12*, 1474–1482.

Wenick, A.S., and Hobert, O. (2004). Genomic cis-Regulatory Architecture and trans-Acting Regulators of a Single Interneuron-Specific Gene Battery in *C. elegans*. Developmental Cell *6*, 757–770.

White, J.G., Southgate, E., Thomson, J.N., and Brenner, S. (1986). The Structure of the Nervous System of the Nematode *Caenorhabditis elegans*. Philosophical Transactions of the Royal Society of London B. Biological Sciences *314*, 1–340.

Wichterle, H., Gifford, D., and Mazzoni, E. (2013). Mapping Neuronal Diversity One Cell at a Time. Science *341*, 726–727.

Williams, D.C., Cai, M., and Clore, G.M. (2004). Molecular Basis for Synergistic Transcriptional Activation by Oct1 and Sox2 Revealed from the Solution Structure of the 42-kDa Oct1{middle dot}Sox2{middle dot}Hoxb1-DNA Ternary Transcription Factor Complex. Journal of Biological Chemistry *279*, 1449–1457.

Wood, W.B. (1988). Introduction to *C. elegans* biology. In The Nematode *Caenorhabditis Elegans*, pp. 1–16.

Wrischnik, L.A., and Kenyon, C.J. (1997). The role of *lin-22*, a hairy/Enhancer of split homolog, in patterning the peripheral nervous system of *C. elegans*. Development *124*, 2875–2888.

Xie, B., Charlton-Perkins, M., McDonald, E., Gebelein, B., and Cook, T. (2007). Senseless functions as a molecular switch for color photoreceptor differentiation in *Drosophila*. Development *134*, 4243–4253.

Xue, D., Finney, M., Ruvkun, G., and Chalfie, M. (1992). Regulation of the *mec-3* gene by the *C. elegans* homeoproteins UNC-86 and MEC-3. The EMBO Journal *11*, 4969–4979.

Xue, D., Tu, Y., and Chalfie, M. (1993). Cooperative interactions between the *Caenorhabditis elegans* homeoproteins UNC-86 and MEC-3. Science *261*, 1324–1328.

Yet, S.-F., McA'Nulty, M.M., Folta, S.C., Yen, H.-W., Yoshizumi, M., Hsieh, C.-M., Layne, M.D., Chin, M.T., Wang, H., Perrella, M.A., et al. (1998). Human EZF, a Krüppel-like Zinc Finger Protein, Is Expressed in Vascular Endothelial Cells and Contains Transcriptional Activation and Repression Domains. Journal of Biological Chemistry *273*, 1026–1031.

Zhang, F., Bhattacharya, A., Nelson, J.C., Abe, N., Gordon, P., Lloret-Fernandez, C., Maicas, M., Flames, N., Mann, R.S., Colón-Ramos, D.A., et al. (2014). The LIM and POU homeobox genes *ttx-3* and *unc-86* act as terminal selectors in distinct cholinergic and serotonergic neuron types. Development *141*, 422–435.

Zhang, P., Behre, G., Pan, J., Iwama, A., Wara-aswapati, N., Radomska, H.S., Auron, P.E., Tenen, D.G., and Sun, Z. (1999). Negative cross-talk between hematopoietic regulators: GATA proteins repress PU.1. Proceedings of the National Academy of Sciences *96*, 8705–8710.

Zhang, P., Zhang, X., Iwama, A., Yu, C., Smith, K.A., Mueller, B.U., Narravula, S., Torbett, B.E., Orkin, S.H., and Tenen, D.G. (2000). PU.1 inhibits GATA-1 function and erythroid differentiation by blocking GATA-1 DNA binding. Blood *96*, 2641–2648.

Zhang, Y., Ma, C., Delohery, T., Nasipak, B., Foat, B.C., Bounoutas, A., Bussemaker, H.J., Kim, S.K., and Chalfie, M. (2002). Identification of genes expressed in *C. elegans* touch receptor neurons. Nature *418*, 331–335.

Zupanc, G.K.H. (1996). Peptidergic transmission: From morphological correlates to functional implications. Micron *27*, 35–91.

CHAPTER 2: MECHANISMS OF HOMEOTIC TRANSFORMATION BETWEEN ALM AND BDU NEURONS

The results shown in this chapter address a number of important issues that remain unresolved in both *C. elegans* and vertebrate systems. First, little is known about the regulation of neuropeptidergic identity. In this section, I present evidence showing that BDU peptidergic identity is regulated by a combination of master regulators and subroutine-specific transcription factors. Second, I examine the means by which cells endogenously repress alternate fate and present a novel competition mechanism for simultaneous activation of one cell identity and repression of another.

I carried out all the experiments described in this chapter. Technical assistance was given by Chi Chen, who did the microinjections.

1. Abstract

Neuron identity transformations occur upon removal of specific regulatory factors in many different contexts in vertebrate or invertebrate nervous systems, thereby revealing the fundamental principle of alternative cell identity choices made during nervous system development. One common molecular interpretation of such homeotic cell identity transformations is that a regulatory factor activates genes defining one cellular identity, while simultaneously repressing the expression of genes that define an alternative identity. We provide here evidence for an alternative mechanistic basis of a homeotic identity transformation. We show that the identity of the neuropeptidergic BDU interneuron is defined by two collaborating terminal selector-type transcription factors, the POU homeobox gene *unc-86* and the Zn finger factor *pag-3*. The identity of the sister cell of BDU, the ALM neuron also employs the *unc-86* gene as a terminal identity selector, but even though *pag-3* is also expressed in ALM, *unc-86* rather cooperates with a distinct transcription factor, the LIM homeobox gene *mec-3*. We show that in the absence of *mec-3*, the ALM neuron homeotically transforms into the BDU neuron, in an *unc-86*- and *pag-3*-dependent manner. MEC-3 prevents the execution of UNC-86/PAG-3-dependent BDU identity by directly interacting with UNC-86 and thereby preventing UNC-86 from engaging with PAG-3 to induce BDU identity. The basic principle that our study reveals is that the promotion of a specific identity program and simultaneous inhibition of an alternative program is not the result of distinct activation and repression capabilities of a transcription factor, but rather caused by the ability of a transcription factor to outcompete the execution of an alternative program. Homeotic control of neuronal identity programs has implications for the evolution of neuronal cell types.

2. Introduction

In 1894, Bateson introduced the term homeosis, or homeotic transformation, to describe transformations of identities of homologous characters in a repeated series of animal characters (e.g. vertebrae). He observed these transformations as naturally occurring variants within many different species (Bateson, 1894). Homeotic transformations are not limited to segmented structures but can refer to different levels of organization, generally describing any transformation of one part of an organism into another (Holmes, 1986; Leavitt, 1909; Sattler, 1988).

In addition to whole tissues or organs, the homeosis concept has been applied to the level of single cells. For example, many classic lineage mutants in the nematode *Caenorhabditis elegans*, such as *lin-4, lin-12/Notch* or *lin-22/Hairy*, cause cellular identity transformations on a single-cell level that have been described as homeotic (Sternberg and Horvitz, 1984). Photoreceptor identity transformations in the *Drosophila* retina, observed for example upon removal of the *sevenless* gene, have also been characterized as homeotic transformations (Tomlinson and Ready, 1986). In vertebrates, removal (or ectopic expression) of specific regulatory factors within many tissue types have also been shown to result in cell identity switches. For example, in the immune system, loss of the transcription factor (TF) PU.1 results in a conversion of myeloid cells to erythroid cells, while loss of GATA1 results in a conversion of erythroid cells to myeloid cells (Graf and Enver, 2009; Iwasaki and Akashi, 2007).

In the context of the nervous system, a variety of studies have shown that loss of expression or ectopic expression of some TF can bring about cell identity switches that

are essentially homeotic in nature (Holmberg and Perlmann, 2012). For example, in mouse striatal interneurons, the LIM homeobox gene *Lhx7* promotes cholinergic fate; loss of *Lhx7* causes those neurons to instead adopt GABAergic fate (Lopes et al., 2012). In the dorsal horn of the spinal cord, *Lbx1* selects GABAergic cell fate over glutamatergic cell fate (Cheng et al., 2005), while in the mesencephalon, *Helt* induces GABAergic fate while repressing glutamatergic fate (Nakatani et al., 2007). Distinct cortical neuron types in different cortical layers switch their identity upon removal of different types of TFs (Srinivasan et al., 2012). In the *C.elegans* ventral nerve cord, VD motor neurons switch their identity to lineally unrelated DD motor neurons upon removal of the TF *unc-55* (Shan et al., 2005). With the caveat in mind that the true extent of transformation in cellular identity can often only be assessed with a limited spectrum of available tools, it nevertheless appears evident that binary cell fate choices, whose disruption results in homeotic cell identity transformations, are a recurring theme in patterning of the nervous system.

The mechanistic basis of transformations in cell identity is often not clear. In principle, a transcription factor can simultaneously operate as an activator for some targets and a repressor of other target genes. In such cases, genetic removal of the TF results in failure to activate gene batteries that define one cellular state and a derepression of gene batteries that define an alternative state. Indeed, it has been shown that in the context of neocortical projection neurons, Fezf2 can activate genes that define the glutamatergic phenotype while directly repressing genes that define the GABAergic phenotype (Lodato et al., 2014). Cross- repressive interactions between TF inducers of specific identity programs have also been observed, for example, in the immune system

(Graf and Enver, 2009) or the nervous system (Srinivasan et al., 2012). In this paper we describe a novel principle that underlies a homeotic neuronal identity transformation in the nervous system of the nematode *C. elegans.*

We have studied a homeotic neuronal identity transformation in the context of two *C. elegans* neuron classes, the ALM and BDU classes . The ALM neuron class is composed of a pair of two bilaterally symmetric neurons (ALM left and ALM right) that sense light touch in the anterior half of the animal, along which they extend axon projections (Chalfie et al., 1985). The sister cells of the left and right ALM neurons, the left and right BDU neurons, are generated during embryonic stages by asymmetric division of a neuroblast (Sulston et al., 1983) and are located in a more anterior lateral midbody domain of the animal. The BDU neurons extend their axons alongside, but in a more ventral position as those of the ALM neurons; the BDU axons then turn ventrally through the anterior deirid commissure and eventually terminate after a dorsal turn into the nerve ring (White et al., 1986).

BDU neurons are distinct from their sister ALM neurons on many different levels. Unlike the ALM neurons, which contain specialized microtubules, or other sensory neurons that are highly branched or contain ciliated endings, the BDU neurons do not show any specific morphological features that would suggest a sensory neuron function (White et al., 1986). However, recent cell ablation studies have demonstrated that the BDU neurons are involved in a harsh touch response to the anterior half of the animals (Li et al., 2011). Whether the BDU neurons are themselves mechanoreceptors or act downstream of a mechanosensory neuron is presently not clear, but the conspicuous extension of BDU processes along the anterior half of the animal suggests that these

neurons may indeed be mechanosensory.

Apart from their striking morphological differences, there are also notable differences in the neurotransmitter choice of the ALM and BDU neurons. The ALM neurons are glutamatergic (Lee et al., 1999). In contrast, unlike most *C.elegans* neurons, all of the synaptic outputs of the BDU neurons contain striking, darkly staining vesicles, suggesting that the BDU neurons make prominent use of neuropeptides (White et al., 1986). Indeed, five neuropeptide-encoding genes, producing at least 11 different neuropeptides are known to be expressed in BDU: *flp-10, flp-12, nlp-1, nlp-15* and *nlp-37* (Kim and Li, 2004; Li and Kim, 2010; Nathoo et al., 2001). Moreover, a systematic mapping of neurotransmitter systems suggests that BDU may not use any classic, fast-acting neurotransmitter system, such as acetylcholine, glutamate, GABA or monoamines (Duerr et al., 2001; McIntire et al., 1993; Serrano-Saiz et al., 2013)(our unpublished data). This indicates that the BDU neuron class may be akin to other neurons in vertebrate and invertebrate nervous systems, such as vertebrate oxytocin/vasopressin-expressing magnocellular neurons which exclusively utilize neuropeptides for communication with downstream target neurons (Salio et al., 2006). How peptidergic neurotransmitter identity of a neuron is controlled and coupled to other identity features of a neuron is not well understood.

While previous work has identified two TFs, the *unc-86* POU homeobox and the *mec-3* LIM homeobox genes, as critical regulators of ALM identity (Duggan et al., 1998; Way and Chalfie, 1988, 1989; Xue et al., 1992; Xue et al., 1993; Zhang et al., 2002), the differentiation program of the BDU neurons has not previously been investigated in depth. So far, it has only been shown that the *C.elegans* Senseles/Gfi ortholog *pag-3* in

repressing aberrant expression of two ALM marker genes in BDU (Jia et al., 1997). We demonstrate here how BDU neuropeptidergic identity is controlled and how the BDU differentiation program relates to the adoption of ALM glutamatergic identity. We find that, like in ALM, the *unc-86* POU homeobox gene is critically required for the establishment of BDU identity. *unc-86* operates with a different factor in BDU than in ALM. While it operates with *mec-3* in ALM, it works together with the Zn finger TF *pag-3* in BDU. We show that loss of either *pag-3* or *mec-3* causes reciprocal homeotic transformations between ALM and BDU. We define a novel mechanism by which *mec-3* is able to both induce ALM fate and repress BDU fate that involves competition for access to the UNC-86 TF.

3. Materials and Methods

3.1. Mutant alleles

unc-86(n846), unc-86(n848), flp-10(ok2624), nlp-15(ok1512), *nlp-1(ok1469), egl-3(nr2090), pag-3(ls20), ahr-1(ia3), ceh-14(ch3), unc-86(u5), unc-86(u168), mec-3(e1338), mec-3(u298), mom-4(ne1539); lit-1(t1512).*

3.2. DNA constructs

All *gfp* constructs for the mutational *cis*-regulatory analysis were generated by subcloning promoter fragments into the pPD95.75 multiple cloning site (Addgene). The *hsp::unc-86* and *hsp::mec-3* constructs were generated by subcloning cDNA into the vector pPD49.78 (Addgene). Linearized plasmids were injected as complex arrays using 5 ng/ul (for *gfp* constructs) or 2 ng/ul (for *hsp* constructs) of plasmid, 100 ng/ul of PvuII-digested bacterial genomic DNA, and 5 ng/ul of injection marker. Mutagenesis was

performed using the QuikChange II XL Site-directed Mutagenesis Kit (Stratagene). Fosmid reporters were recombineered as previously described (Tursun et al., 2009), using fosmid WRM0612cF07 for *unc-86* and WRM0632dA05 for *pag-3*. For the *unc-86* fosmid reporter constructs, the *yfp* and *mChOpti* coding sequences were added at the C-terminal end of the protein after the *gpd-2* intergenic region (which contains an SL2 trans splicing signal), while for the *pag-3* fosmid, the reporter was added at the N-terminal end. The *unc-86* fosmids were injected into *unc-86(n846)*, while the *pag-3* fosmids were injected into *pag-3(ls20)*. Fosmid reporters were functional as assessed by rescue of mechanosensory defects for *unc-86* and by rescue of the reverse kinker unc phenotype for *pag-3.* Linearized fosmids were injected at 10 ng/ul with 100 ng/ul of PvuII-digested bacterial genomic DNA and 5 ng/ul of injection marker (*ttx-3::mCherry* for *unc-86* and *ttx-3::gfp* for *pag-3*). *ceh-14* and *mec-3* fosmid reporters were obtained from the ModEncode consortium.

4.3. Harsh touch assay

Harsh touch assays were performed on gravid adults which were transferred to a Nematode Growth Medium (NGM) plate with OP50 bacteria one hour before testing. Animals were scored blind to genotype. The assay was performed on animals moving in a forward direction while off the bacteria lawn. Each animal was touched once in the anterior half of the body, just posterior to the pharynx, using a flattened platinum wire pick attached to a glass Pasteur pipette. Animals were scored by measuring the number of head swings each animal moved in a backwards direction before stopping, reversing direction, or performing an omega turn.

3.4. Heat shock experiments

For *hsp::mec-3* experiments, animals were heat shocked at bean stage (~360 minutes), 3-fold stage (~550 minutes), or at L2 larval stage. For *hsp::unc-86* experiments, animals were heat shocked at 3-fold stage. Each heat shock consisted of three rounds of 30 minutes at 37°C, followed by one hour of rest at 20°C. Animals were then maintained overnight at 25°C and scored the following day.

3.5. Electrophoretic mobility shifts

Full-length *mec-3* cDNA was cloned into the pET21b His tag bacterial expression vector (EMD Millipore) and transformed into BL21(DE3) cells (NEB). Protein expression was induced using 1 mM IPTG for four hours and purified with Ni-NTA resin (Qiagen) under denaturing conditions as previously described (Wenick and Hobert, 2004). *unc-86* cDNA in pET21b was similarly purified, as previously described (Zhang et al., 2014). *ceh-43* cDNA in pET21b was induced in 1 mM IPTG for four hours. To purify, bacteria was pelleted, frozen, and resuspended in 50 mM Tris (pH 7.5), 500 mM NaCl, 20 mM Imidazole with protease inhibitors. The solution was sonicated and purified with Ni-NTA resin (Qiagen). The same buffer plus 300 mM imidazole was used for elution, and protein was dialyzed into 20 mM HEPES (pH 7.9), 200 mM NaCl, 10% glycerol, 2 mM $MgCl_2$.

To perform EMSAs, a short oligonucleotide was end-labeled with [γ-^{32}P]ATP using T4 polynucleotide kinase (NEB) according to the manufacturer's specifications. A complementary sequence was added to the 3' end of each oligonucleotide used in the

EMSA. The radiolabeled short sequence was annealed to each long oligo and the remaining DNA was filled in using Klenow (NEB). Protein and DNA were incubated for 20 minutes at room temperature in 1x binding buffer (5x Binding Buffer (BB): 50 µl 1M Tris pH 7.5, 50 µl 5M NaCl, 5 µl 1M $MgCl_2$, 250 ul 80% glycerol, 2.5 µl 1 M DTT, 5 µl 0.5M EDTA, 250 µl Poly dI-dC, 2.5mg/ml bovine serum albumin, 290 µl H_2O), before loading on 4% (79:1 acrylamide:bis-acrylamide) gel and run at 165V at 4°C for 2-3 hours. Purified UNC-86 was run at 100 nM concentration. MEC-3 concentrations for the *ceh-14* probe were 50, 100, and 200 nM; for the *tph-1* probe, MEC-3 concentrations were 100 and 200 nM. CEH-43 concentrations were 50, 100, and 200 nM.

The EMSA probe sequences are as follows (underlining indicates the sequence added for complementarity to the short labeled oligonucleotide):

ceh-14 probe:

AATTGTTTTCATTTAAAATGAGCAACTGTAATTTTCTATTCATTAAAGTATTTT
TTTTACCATTTAAAAGGAACCCATTCATGAAAAGTTGGTCAAGGTCGTTTCC

tph-1 probe:

CCCAACACCACATTATTCATGTATTTCCTCCAAACCACTGAACCATCTCATTC
TCAAACCAGTTTCTATCCGTTTGTTTGCATTCAATTAAATTTTTGGTCAAGGTC
GTTTCC

unlabeled *mec-3*:

ACATTTGAAAAAACAACAAATTCATTCGAAATGCATTGCCCATAATGAATCG

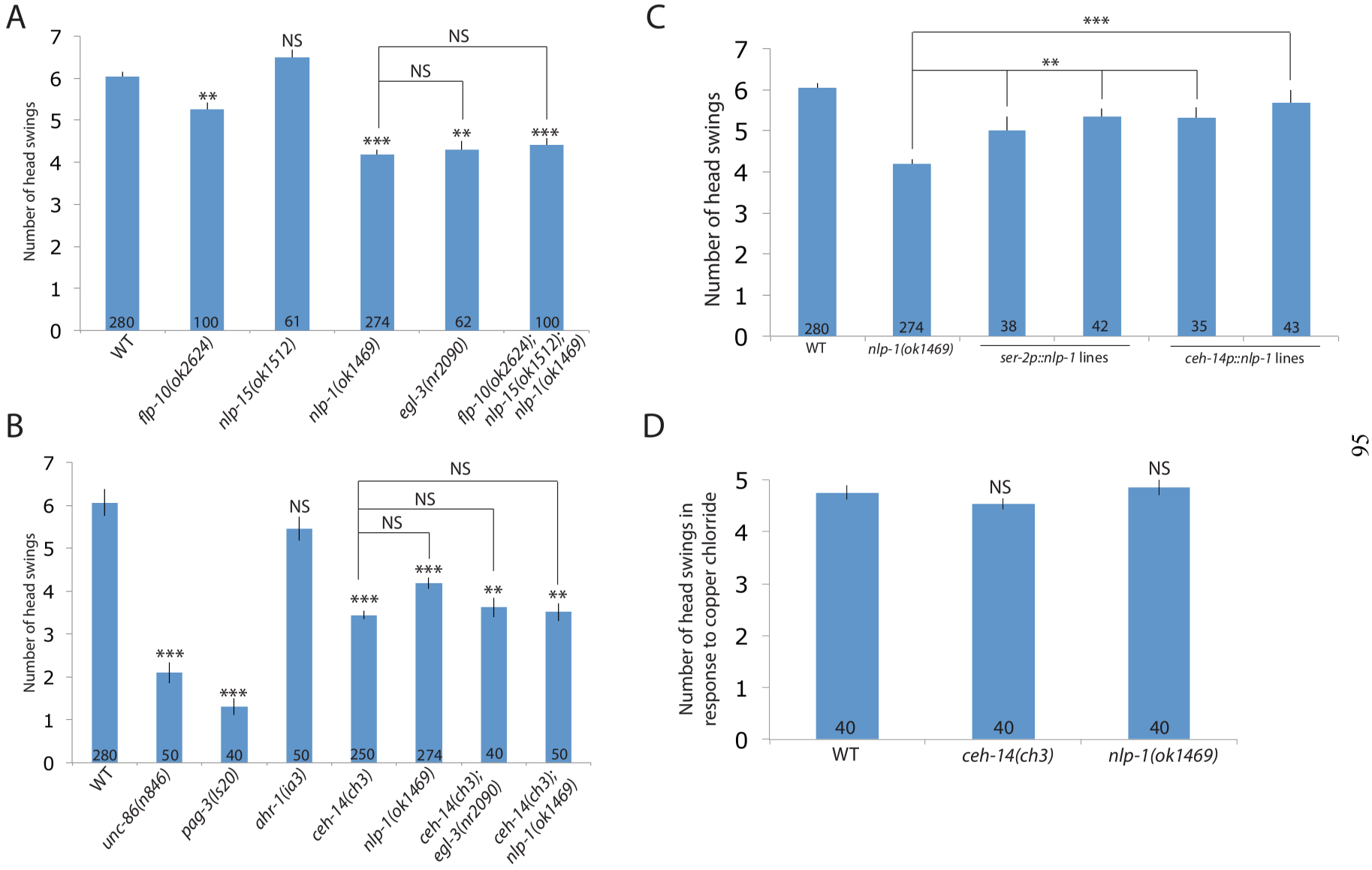
ACCGAAAAACACAAGTGACCGTCAGGAGATCGATAGAG

4. Results

4.1. Neuropeptides mediate the function of BDU in touch sensation

Our first goal was to understand the genetic basis for BDU function. As mentioned in the introductory section, the BDU neurons are peptidergic, expressing at least eleven distinct neuropeptides, encoded by 5 different neuropeptide encoding genes (*flp-10, flp-12, nlp-1, nlp-15* and *nlp-37*) (Kim and Li, 2004; Li and Kim, 2010; Nathoo et al., 2001). We sought to test whether the known mechanosensory function of the BDU neurons indeed requires neuropeptide-mediated signaling by examining animals that carry null mutant alleles of the *flp-10, nlp-1* and *nlp-15* loci, respectively. We found that *flp-10* and *nlp-1* single mutants show defects in harsh touch responses, while *nlp-15* null mutants do not (Figure 1A) We focused on the *nlp-1* mutant defects because they are the strongest. *nlp-1* mutants can also not be further enhanced by simultaneous removal of *flp-10* or *nlp-15* (Figure 1A). Moreover, the *nlp-1* harsh touch defects are not due to a general unresponsiveness to engage in an "escape reflex", as *nlp-1* mutants still are repelled by noxious copper sulfate (Figure 1D). To ascertain the focus of action of *nlp-1*, we generated transgenic animals expressing *nlp-1* under the control of two different, BDU-expressed drivers, one from the *ser-2* locus, the other from the *ceh-14* locus. Each driver is expressed in a small number of additional neurons, but they only overlap in their expression in BDU. We found that *nlp-1* expressed by either driver is able to rescue the *nlp-1* null mutant defect (Figure 1C).

Figure 1: BDU-mediated mechanosensory responses are mediated by BDU expressed neuropeptides



Gravid adults were touched with a platinum pick in the anterior half of the midbody. The number of head swings of backwards movement before animals stopped, reversed direction, or did an omega turn was scored. *n* is given at the bottom of each bar. NS, not significant **p<.001 ***p<.0001 (t-test with Bonferroni correction). Error bars: s.e.m.

(A) Mechanosensory response of wildtype (WT), neuropeptide mutants, and neuropeptide processing mutants.

(B) Mechanosensory response of WT and transcription factor mutants.

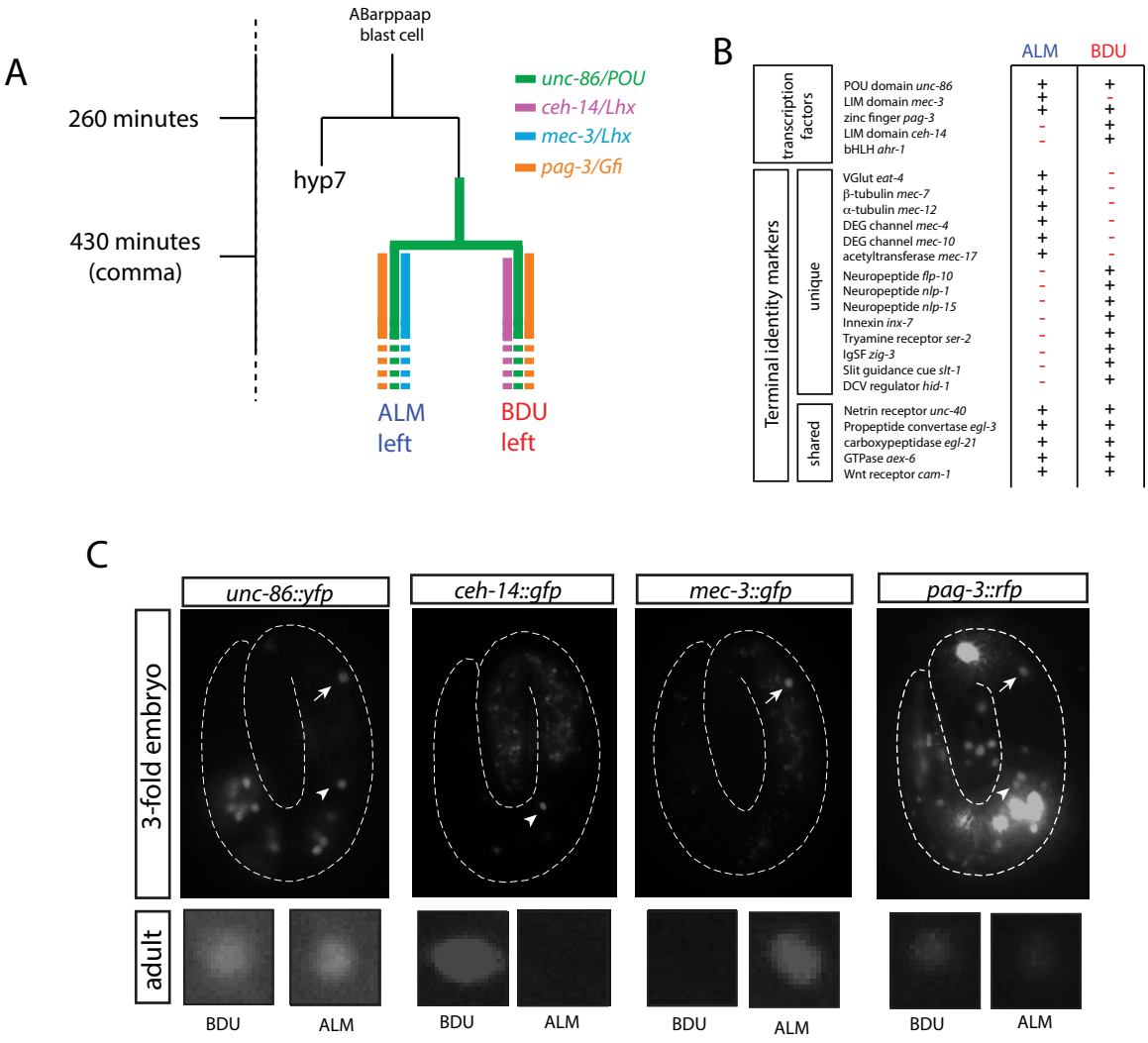
(C) Rescue of the *nlp-1(ok1469)* phenotype by expressing *nlp-1* under control of two different BDU promoters.

(D) To ensure that *ceh-14(ch3)* and *nlp-1(ok1469)* mutants are not generally defective in backwards response, animals were exposed to the noxious chemical copper chloride. No difference was seen between wild type and *ceh-14* or *nlp-1* animals.

4.2. *ceh-14* controls the peptidergic identity of BDU neurons

While numerous studies have elucidated the mechanisms that genetically control classic neurotransmitter identity (Hobert, 2011), little is known about the genetic control and coordination of peptidergic neuron identity features. We sought to address this question by considering the function of transcription factors previously shown to be expressed in the BDU neurons: the *unc-86* POU homeobox gene (Finney and Ruvkun, 1990), the *pag-3* gene, an ortholog of the Senseless/Gfi Zn finger TF (Jia et al., 1997), and the *ceh-14* LIM homeobox gene, the ortholog of vertebrate Lhx3/4 (Cassata et al., 2000) (Figure 2A**)**. We confirmed these expression patterns using fosmid-based reporters containing approximately 40kb of genomic sequence. We found that each TF is expressed throughout the life of the BDU neurons (Figure 2C). The *unc-86* gene is turned on in the mother of the BDU neuron (Finney and Ruvkun, 1990), while *ceh-14* and *pag-3* start to be expressed after the birth of the BDU neurons (Figure 2C, summarized in Figure 2A). Consistent with its earlier onset of expression, *unc-86* is required for expression of *pag-3* and *ceh-14* (Figure 3A).

Figure 2: ALM and BDU transcription factor expression

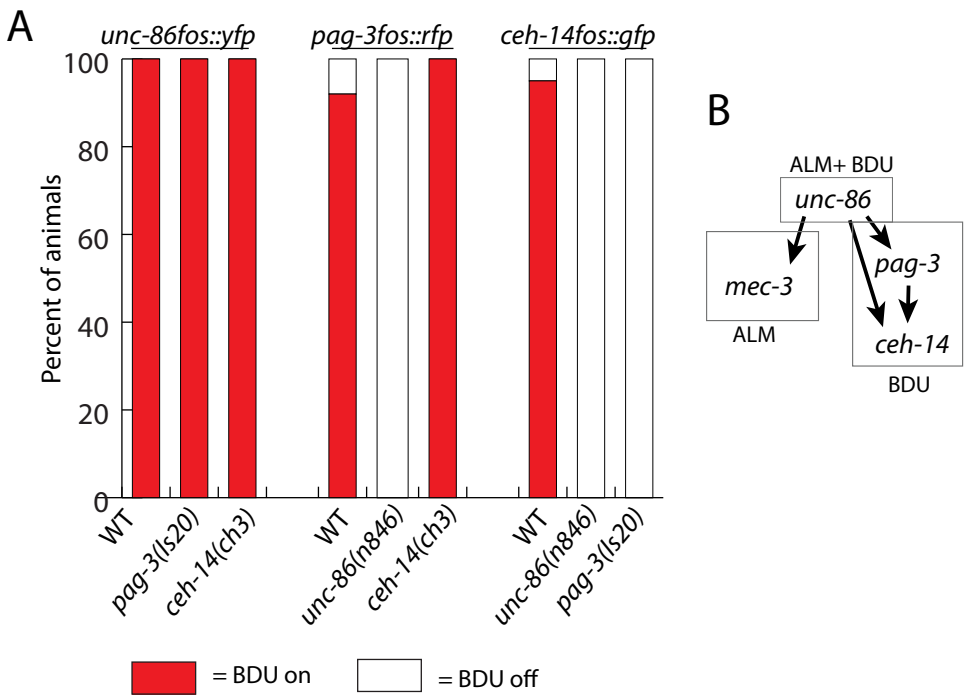


(A) Timing of the ALM/BDU cell division and transcription factor expression. The ALM/BDU mother divides around 430 minutes post fertilization. *unc-86* is expressed beginning in the ALM/BDU mother cell, while *mec-3*, *pag-3*, and *ceh-14* are expressed from shortly after the ALM/BDU cell division. Dotted lines indicate continuous expression in all adult stages.

(B) Lineage & terminal markers of ALM and BDU. '+' and '-' indicate expression of each gene in the given neuron

(C) Expression of fosmid-based reporters of transcription factors in embryo and adult. *unc-86*^{fosmid}*::yfp* (*otIs337*) is expressed in ALM and BDU. *ceh-14*^{fosmid}*::gfp* (*wgIs73*) is expressed in BDU. *mec-3* (*wgIs55*) is expressed exclusively expressed in ALM. *pag-3*^{fosmid}*::rfp* (*otIs429*) is expressed in ALM and BDU. Arrows in embryo images indicate ALM, arrowheads indicate BDU.

Figure 3: ALM and BDU transcription factor regulatory hierarchy

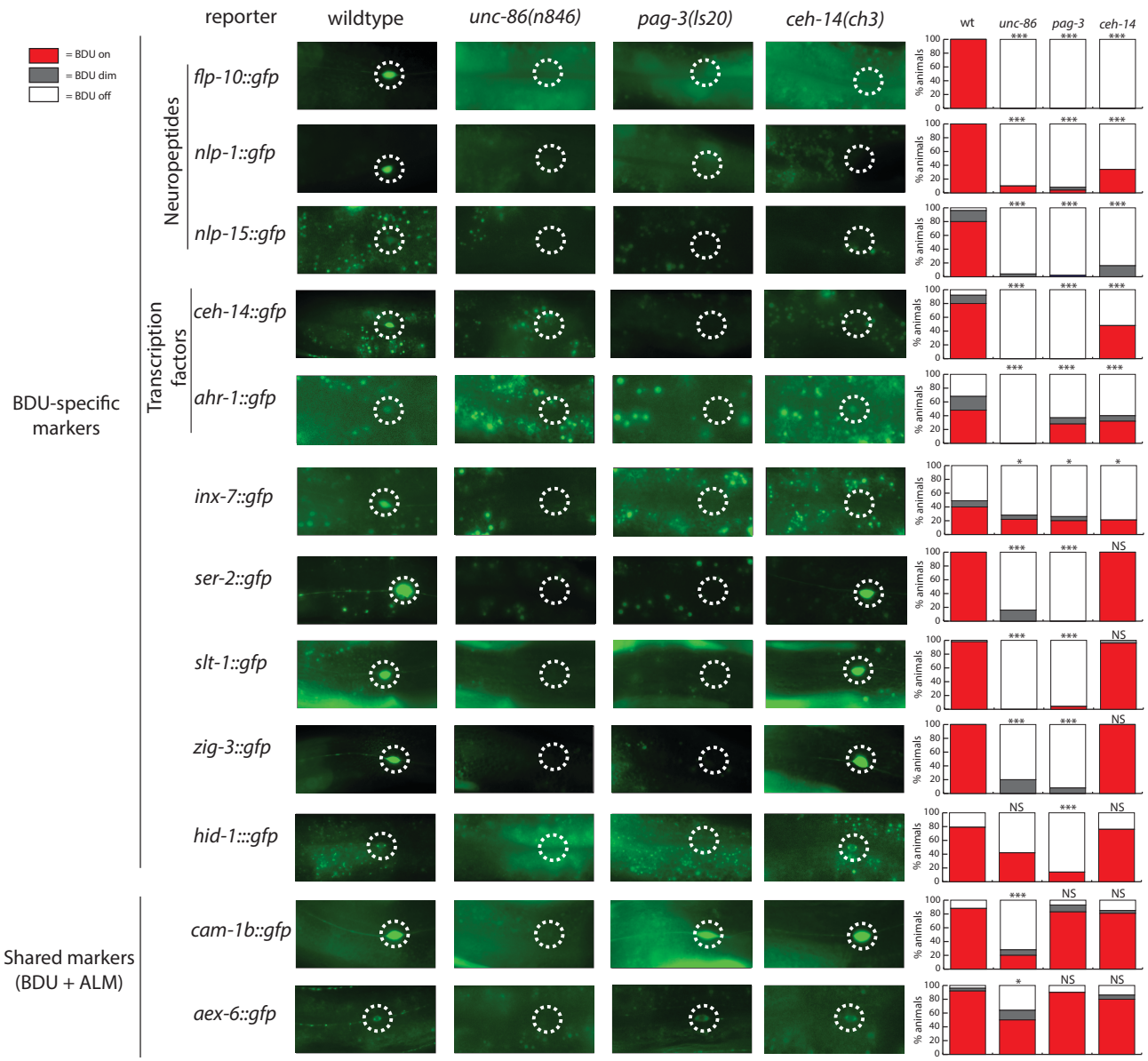


(A) Cross-regulation of BDU transcription factors. While expression of *unc-86* is unaffected by either *pag-3(ls20)* or *ceh-14(ch3)*, expression of *pag-3* is off in BDU in *unc-86(n846)* and expression of *ceh-14* is off in both *unc-86(n846)* and *pag-3(ls20)*. Animals were scored at the L4 stage. $n \geq 50$.

(B) Summary of transcription factor interactions. Regulation of *ceh-14* by both *pag-3* and *unc-86* ("feedforward loop") is inferred from the analysis of the *cis*-regulatory architecture of the *ceh-14* locus described in Fig.6B. *mec-3* regulation by *unc-86* was previously shown (Xue et al., 1992).

The availability of eight molecular markers of terminal BDU identity (listed in Figure 2B) allowed us to probe the effect of BDU-expressed TFs on terminal BDU identity. We found that *ceh-14* null mutant animals lose the expression of only a subset of the BDU identity markers, including of all three neuropeptide genes examined, *flp-10, nlp-1* and *nlp-15* (Figure 4). As expected from the loss of *nlp-1* expression, we found that *ceh-14* mutants display harsh touch response defects (Figure 1B). These defects are not further enhanced by removal of either *nlp-1* or a gene*, egl-3,* that is generally required for neuropeptide processing (Figure 1B), indicating that the defects in *ceh-14* mutants can be ascribed to their loss of neuropeptide signaling. *ceh-14* also controls its own expression (Figure 4), suggesting that *ceh-14* may be continuously required to maintain neuropeptidergic identity. The expression of *unc-86* and *pag-3* are unaffected in *ceh-14* null mutants (Figure 3A).

Figure 4: The effect of *unc-86*, *pag-3*, and *ceh-14* on BDU identity

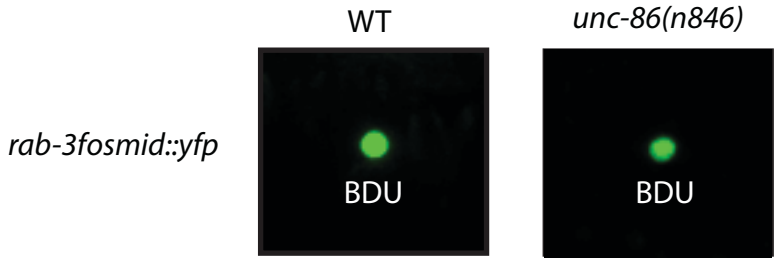


Reporters for BDU genes are shown in wild-type (WT), *unc-86(n846)*, *pag-3(ls20)*, and *ceh-14(ch3)* backgrounds. Right: quantification of expression. $n \geq 50$ for each reporter. See Supplementary Material for information on reporter transgenes. Significance is indicated in relation to wild type tested with Fischer's Exact Test. ***:$p \leq .0001$, *:$p \leq .05$, NS: not significant

4.3. *unc-86* and *pag-3* affect all aspects of BDU identity

In contrast to the restricted defects of *ceh-14* null mutants, loss of the *unc-86* POU homeobox affects all neuron-type specific features of BDU terminal identity (Figure 4). Expression of all neuropeptide-encoding genes is abrogated in *unc-86* mutants; moreover, expression of the INX-7 innexin, the dense core vesicle regulator HID-1, the GTPase AEX-6, the tyramine receptor SER-2, secreted proteins (SLT-1, ZIG-3), the CEH-14 and PAG-3 transcription factors and the transcription factor AHR-1 and its target CAM-1 are also affected in *unc-86* (Figure 4). *unc-86* mutants also display defects in the harsh touch response (Figure 1B). To assess whether *unc-86* is required for the generation of the BDU neurons, we examined the expression of the panneuronal marker *rab-3* in *unc-86* mutants and found it to be normally expressed (Figure 5). We conclude that *unc-86* is not required for the generation of BDU or adoption of its generic neuronal identity, but it is required for all known aspects of its neuron-type specific terminal differentiation program.

Figure 5: a panneuronal marker is unaffected in *unc-86* mutants



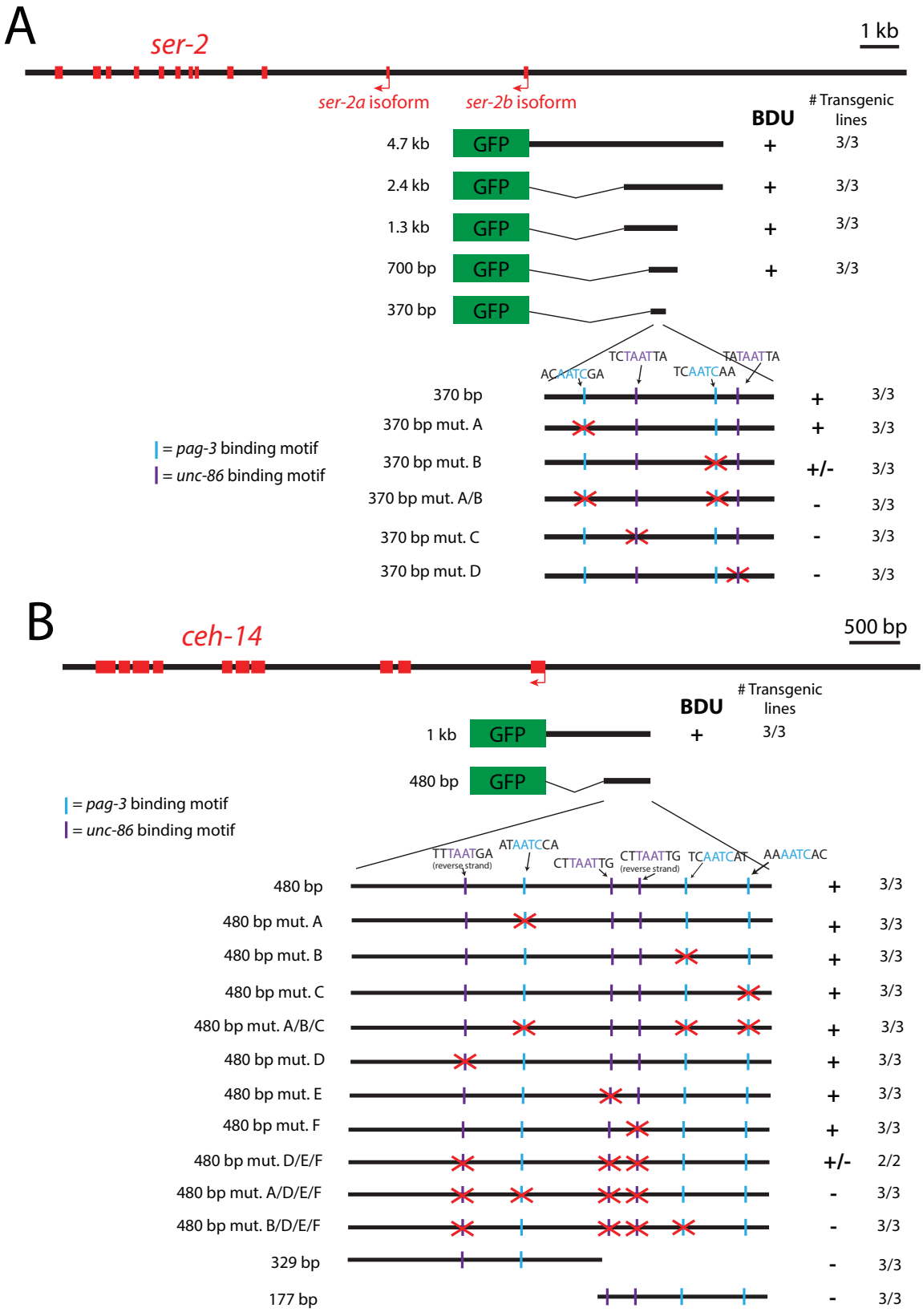
A broadly expressed neuronal gene, *rab-3*, was expressed in both wild type and *unc-86(n846)*. *rab-3::yfp* expression was still seen in BDU in *unc-86*, showing that the BDU neurons are still present.

pag-3 also has a very broad effect on BDU identity. Expression of all of the terminal BDU identity genes described above (except those that are also expressed in ALM) is strongly affected in *pag-3* mutants, including the expression of *ceh-14* and its neuropeptide targets (Figure 4). However, *pag-3* does not have an impact on *unc-86* expression in BDU (Figure 3A). The latter findings demonstrate that in *pag-3* mutants, the BDU neuron is formed and expresses *unc-86*, but is not able to induce terminal BDU differentiation.

To examine whether *unc-86* and *pag-3* directly control terminal BDU identity features, we analyzed the *cis*-regulatory architecture of two genes that define terminal BDU identity and that are both *unc-86*- and *pag-3*-dependent, the tyramine receptor-encoding *ser-2* locus and *ceh-14* LIM homeobox gene. Transgenic reporter animals that contain 5' regions of the *ser-2* and *ceh-14* loci showed expression in BDU and other neurons. We narrowed down these 5' reporters to ~ 400 bp fragments that still yielded expression in the BDU neurons (Figure 6) and examined these fragments for the existence of biochemically defined binding sites of vertebrate and/or fly orthologs of UNC-86 and PAG-3/Senseless. Each of these reporters indeed contains predicted UNC-86/POU binding sites and PAG-3/Senseless binding motifs (Figure 6). We deleted these motifs and found that they are required for expression of *ser-2* and *ceh-14* in BDU. In the case of *ser-2*, deletion of one PAG-3 binding site has intermediate effects on *ser-2* reporter gene expression, while deletion of both predicted PAG-3 binding sites abolishes expression (Figure 6A). Deletion of either POU homeodomain binding site alone abolished *ser-2* reporter expression (Figure 6A). In the case of the *cis*-regulatory controls regions of the *ceh-14* locus, we found three predicted binding sites for either transcription

factor and observed a synergistic requirement for these sites. Deletion of single UNC-86 binding sites had no effect on reporter expression, while deletion of all UNC-86 sites partially disrupted expression (Figure 6B). Combining the mutation of all three UNC-86 sites with a mutation in a presumptive PAG-3 site, which alone has no effect on reporter expression, completely abolished reporter expression (Figure 6B).

Figure 6: *Cis*-regulatory elements controlling BDU expressed genes



(A) Mutational analysis of *cis*-regulatory elements controlling expression of the *ser-2* locus.

(B) Mutational analysis of *cis*-regulatory elements controlling expression of the *ceh-14*.

Blue bar: Predicted PAG-3/Senseless/Gfi binding motif (Zweidler-Mckay et al., 1996). Purple bar: predicted POU domain binding sites (Cartharius et al., 2005). Mutation of each site is indicated by a red cross over the bar. For the case of *pag-3* sites, the central AATC was mutated to CCCC. For the case of *unc-86* sites, the central TAAT was mutated to CCAT. (+): at least 80% of animals showed bright expression in BDU. (+/-): between 10 and 30% of animals showed dim expression in BDU. (-): less than 10% of animals showed dim expression in BDU. $n \geq 50$ for each line.

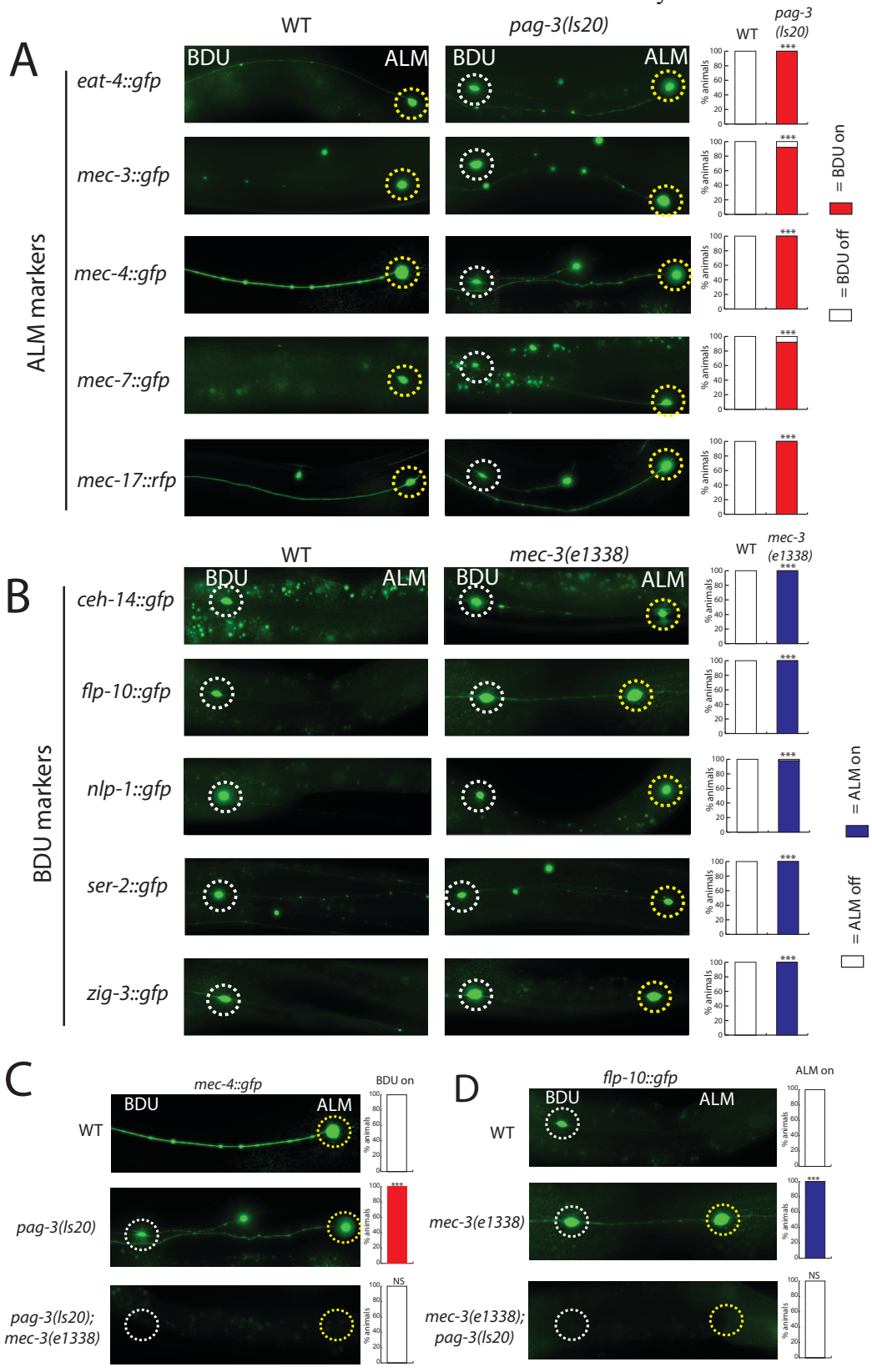
The finding that *ser-2* and *ceh-14* contain functionally required UNC-86 and PAG-3 binding sites suggests that *unc-86* does not merely work through *pag-3* to affect BDU gene expression. We rather conclude that *unc-86* and *pag-3* are terminal selector-type transcription factors (Hobert, 2011) that cooperate to induce terminal differentiation of the BDU neurons.

4.4. Transformation of BDU to ALM identity in *pag-3* mutants

In spite of their similar effects on the induction of terminal features of BDU identity, there are striking differences in the *unc-86* and *pag-3* mutant phenotypes. In the initial identification of *pag-3* mutants, it was noted that the BDU neurons display ectopic expression of two terminal markers of the identity of the ALM sister cell, the *mec-4* ion-channel-encoding gene and the *mec-7*-tubulin-expressing gene (Jia et al., 1996). We examined a potential transformation of BDU to ALM identity in more detail. First, we reexamined the ectopic expression of *mec-4* and *mec-7* in BDU results with different reporters and observed the same results as previously reported (Figure 7A). Second, we examined additional molecular markers for ALM identity, namely the tubulin acetyltransferase-encoding *mec-17* gene (Akella et al., 2010) and the vesicular glutamate transporter-encoding gene *eat-4*, which defines the glutamatergic neurotransmitter identity of ALM (Lee et al., 1999). Both markers are also ectopically expressed in the BDU neurons of *pag-3* mutants (Figure 7A). Third, we examined whether the MEC-4 mechanosensory channel, which normally is targeted to discrete dots along the length of the ALM neuron (Chelur et al., 2002), will cluster along the axon of the transformed BDU neuron and found this to indeed be the case (Figure 8). Fourth, we examined axonal morphology of the BDU neurons in *pag-3* mutants and found that they lose their long

posteriorly directed processes and rather extend short posterior processes much like the ALM neurons (Figure 9A)**.** The ventral turn of the BDU axons into the deirid commissure, normally undertaken by BDU but not ALM, is still executed normally in *pag-3* mutants (Figure 9D) and the axon remains associated with the excretory canal (Figure 9C). We conclude that BDU identity is largely, but not completely transformed into ALM identity in *pag-3* mutants.

Figure 7: Reciprocal homeotic transformation of BDU to ALM in *pag-3* mutants and ALM to BDU in *mec-3* mutants as assessed with molecular identity markers



BDU expression marked with white circles, ALM expression marked with yellow circles. $n \geq 50$. Significance in relation to wild type tested with Fischer's Exact Test. ***:$p \leq .0001$, NS: not significant.

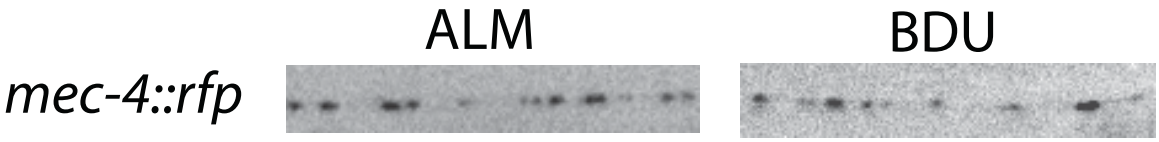
(A) ALM markers are ectopically expressed in BDU in *pag-3(ls20)*.

(B) BDU markers are ectopically expressed in ALM in *mec-3(e1338)*.

(C) Ectopic ALM marker expression is genetically dependent on *mec-3*. Expression of *mec-4^{prom}::gfp* is eliminated in both ALM and BDU in a *pag-3(ls20); mec-3(e1338)* double mutant. The single mutant data is reiterated from panel A for comparison.

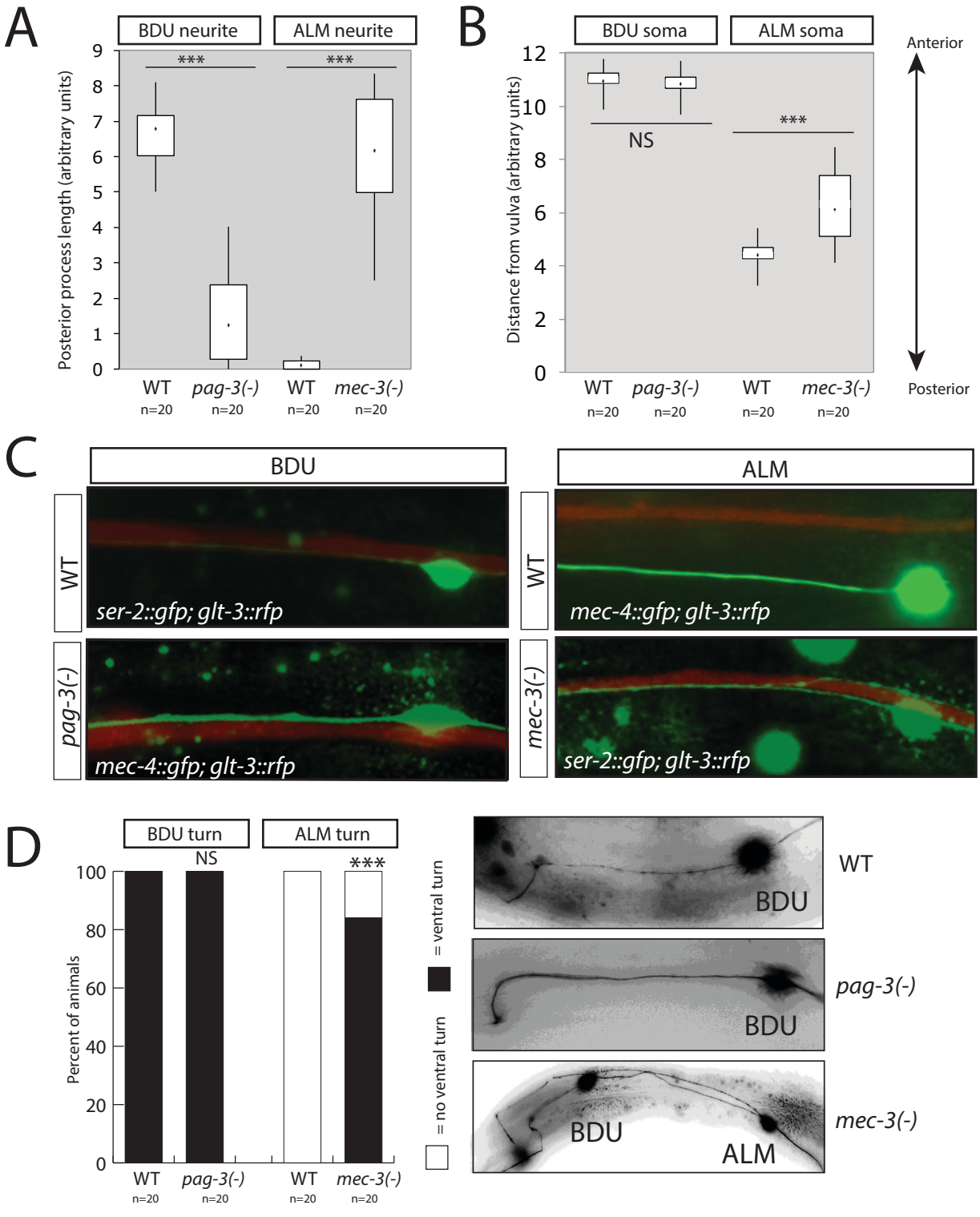
(D) Ectopic BDU marker expression is genetically dependent on *pag-3*. Expression of *flp-10^{prom}::gfp* is eliminated in both ALM and BDU in a *pag-3(ls20); mec-3(e1338)* double mutant. The single mutant data is reiterated from panel B for comparison.

Figure 8: Transformed BDU shows MEC-4 receptor cluster



A *mec-4::rfp* translational fusion appears in puncta along the axon of both ALM and BDU in *pag-3(ls20)*.

Figure 9: Homeotic transformation as assessed by neuronal morphology



(A) The length of the posterior processes of ALM and BDU are altered in *pag-3(ls20)* and *mec-3(e1338)* mutants. Measurements were done on *ser-2::gfp (otIs358)* and *mec-4::gfp* (*zdIs5*)-expressing animals at the L4 stage.

(B) The cell body position of ALM is more anterior in *mec-3(e1338)*, while the position of BDU is unaffected by *pag-3(ls20)*. Measurements refer to distance from the vulva during L4 stage, with higher measurements indicating a more anterior position.

(C) The ventral turn of the BDU process is unaffected in *pag-3(ls20)*, while *mec-3(e1338)* causes the ALM process to undergo a ventral turn. Reporter for wild type and *mec-3(e1338)* is *ser-2::gfp* (*otIs358*); reporter for *pag-3(ls20)* is *mec-4::gfp* (*zdIs5*). The image shown for *mec-3(e1338)* is not typical: in most animals, the ALM axon closely follows the BDU axon in the ventral turn.

(D) ALM is located more laterally in *mec-3(e1338)*, while BDU lateral position is unaffected. Position was examined in relation to the excretory canal (red, *glt-3::rfp)*. The ALM cell body and axon are located more centrally in wild type but colocalize with the excretory canal in *mec-3(e1338)*. BDU is associated with the excretory tract in both wild type and *pag-3(ls20)*.

In all panels, significance refers to comparison to wild type. Significance for A,B measured using a Student's T-test. Significance for C measured using Fisher's Exact Test. NS: not significant ***p<.0001

Previous work had shown that the ectopic expression of *mec-4* and *mec-7* in the BDU neurons of *pag-3* mutants genetically requires *mec-3*, the LIM homeobox gene expressed in ALM and required for ALM differentiation (Jia et al., 1996). We independently confirmed this epistatic relationship with our reporter reagents (Figure 7C). However, the previous *pag-3* study did not provide evidence for ectopic expression of *mec-3* in the BDU neurons of *pag-3* mutants (Jia et al., 1996), as might be expected given the *mec-3* dependence of the BDU to ALM transformation. Using a *mec-3* reporter transgene not previously available, we do observe ectopic expression of *mec-3* in the "BDU" neurons of *pag-3* mutants (Figure 7A). Expression of *mec-3* in ALM is unaffected in *pag-3* mutants (Figure 7A). We conclude that *pag-3* not only drives BDU terminal identity but also represses ALM identity by repressing expression of the ALM identity driver *mec-3* in BDU.

To assess whether the activating effect of *pag-3* on BDU identity could be solely explained by a double-inhibitory mechanism in which *pag-3* inhibits a repressor effect of *mec-3* on BDU identity genes, we analyzed BDU identity in *pag-3; mec-3* double mutants, using the *flp-10* marker (Figure 7D). In these animals, BDU identity is still lost. Therefore, as already suggested by our *cis*-regulatory analysis described above (Figure 6), *pag-3* rather appears to positively induce expression of BDU markers and independently, through repression of *mec-3*, inhibit the expression of ALM identity.

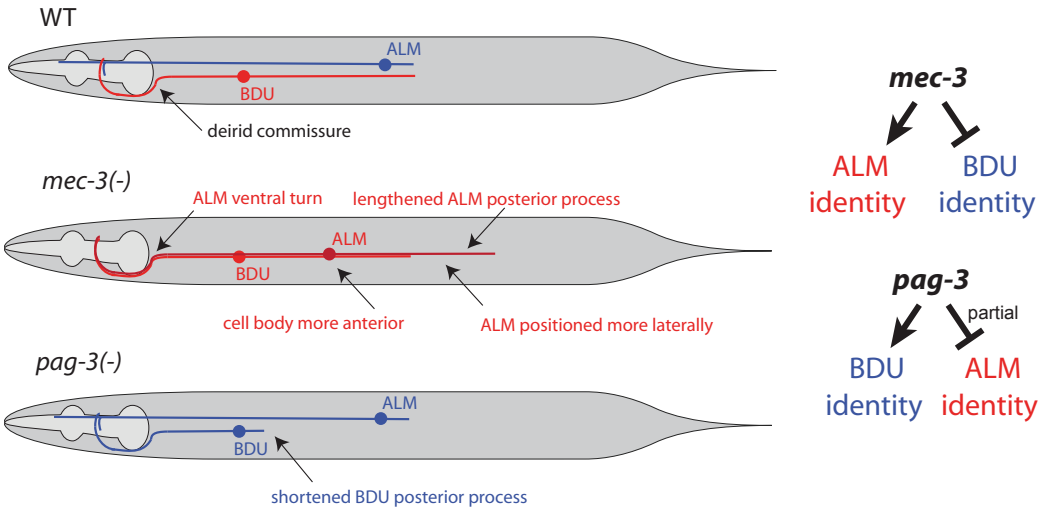
While the loss of BDU identity is shared by *pag-3* and *unc-86* mutants, the BDU to ALM transformation is only observed in *pag-3* mutants and not in *unc-86* mutants. In these animals, no ectopic expression of ALM identity markers can be observed in the

BDU neurons (data not shown). This is expected since ALM differentiation requires *unc-86* (Chalfie et al., 1981; Duggan et al., 1998).

4.5. Reciprocal, homeotic transformation of ALM to BDU in *mec-3* mutants

Previous work has shown that *unc-86* cooperates with *mec-3* to induce terminal differentiation of the ALM neurons (Chalfie et al., 1981; Duggan et al., 1998; Way and Chalfie, 1989; 1988; Xue et al., 1992; 1993). In both *unc-86* and *mec-3* null mutants ALM-specific genes fail to be expressed. Even though the effect of *unc-86* and *mec-3* on the induction of ALM features are similar, there are again striking differences in the *unc-86* and *mec-3* mutants. It was already previously noticed that the axons of the ALM neurons of *mec-3* mutants extend more posteriorly than in wild-type animals and appear more medially and deeper positioned, thereby appearing more BDU-like (Way and Chalfie, 1988). A more anterior position of the cell body of ALM was also noted (Way and Chalfie, 1988). Using previously unavailable *gfp* markers that label BDU morphology, we could examine morphology in greater detail. We confirmed and quantified the presence of the posteriorly directed process of ALM in *mec-3* mutants (Figure 9A), the anterior position of the cell body (Figure 9B) and the ventral shift of the ALM axons in *mec-3* mutants (Way and Chalfie, 1988) (Figure 9C). By colabeling the excretory canal, we found that the transformed ALM axon now occupies the same tract along the excretory canal that the normal BDU axon occupies (Figure 9C). Moreover, we found that the ALM axons of *mec-3* mutants now undergo the ventral turn into the deirid commissure, much alike what BDU axons do (Figure 9D). The morphology transformations are summarized in Figure 10.

Figure 10: Schematic of homeotic transformation



Changes in gene expression in *mec-3*(*e1338*) and *pag-3*(*ls20*) are represented by color, with blue indicating genes required for BDU identity and red indicating genes required for ALM identity. Morphology changes are noted.

The availability of BDU identity markers allowed us to further examine the extent of transformation of ALM to BDU. We found that all BDU identity markers examined, including two neuropeptide-encoding genes, the tyramine receptor *ser-2* and the IgSF *zig-3*, are ectopically expressed in the ALM neurons of *mec-3* mutants (Figure 7B). The ectopic expression of BDU markers in ALM in *mec-3* mutants genetically depends on *pag-3*, since in *mec-3; pag-3* double mutants, the *flp-10* gene fails to be expressed in ALM (Figure 7D).

Taken together, we have shown that in *mec-3* mutants, the ALM neurons display a homeotic transformation to the identity of the BDU neurons, based on morphology (summarized in Figure 10) and molecular profile. Unlike in *pag-3* mutants, where only one morphological feature was transformed, the ALM to BDU transformation in *mec-3* mutants extends to all morphological features that we could examine.

4.6. *mec-3* is restricted to ALM by transcriptional repression in BDU via *pag-3* and a non-canonical Wnt signaling system

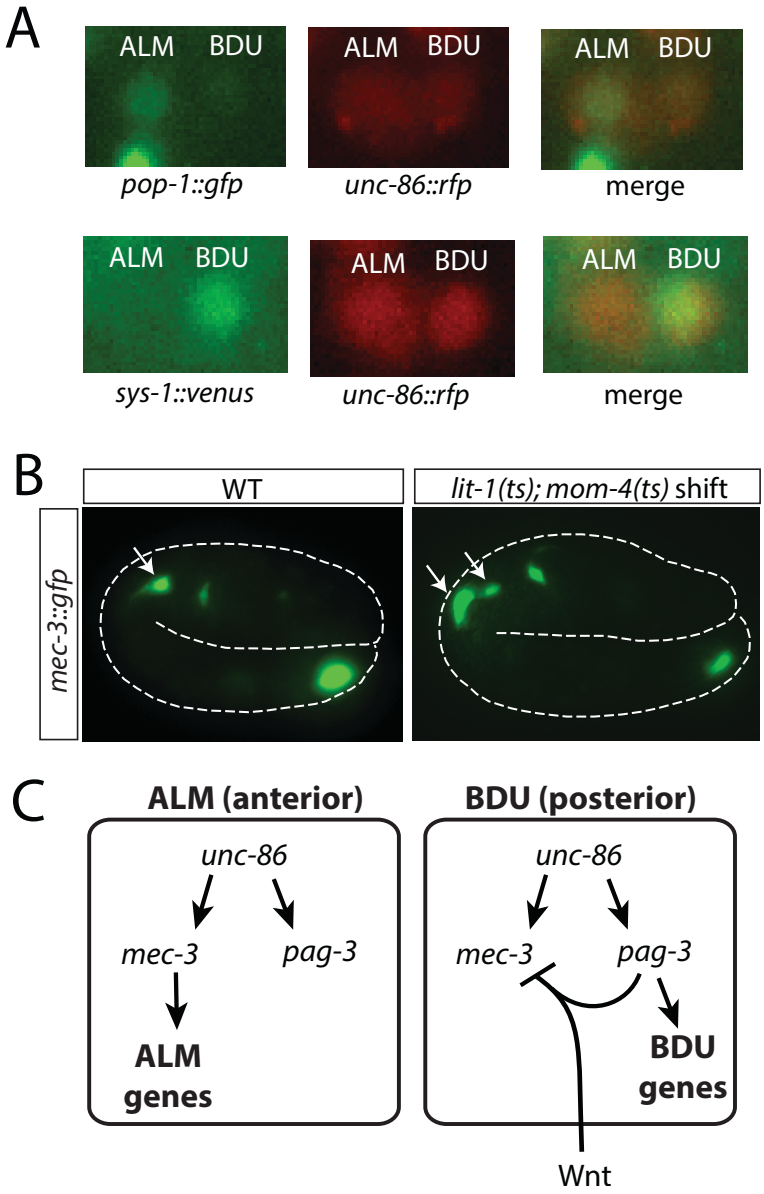
The reciprocal homeotic transformations of neuronal identity in *mec-3* and *pag-3* mutants and the genetic epistasis experiments that we described above demonstrate that *mec-3* and *pag-3* antagonize each others activity. *mec-3* promotes ALM identity and, by antagonizing *pag-3* activity, inhibits BDU identity (i.e. the gain of BDU identity in ALM of *mec-3* mutants genetically depends on *pag-3*), while *pag-3* promotes ALM identity and, by antagonizing *mec-3* activity, inhibits ALM identity (i.e. the gain of ALM identity in BDU of *pag-3* mutants genetically depends on *mec-3*). In principle, such mutual antagonism could occur on the level of a mutual inhibition of each others expression. As described above, we indeed find that *pag-3* represses *mec-3* expression in BDU (Figure

7A). However, both PAG-3 antibody staining (Cameron et al., 2002) and a fosmid-based reporter for *pag-3* expression that we generated (Figure 2C) shows *pag-3* expression not only in BDU, but also in ALM. Expression levels in ALM and BDU appear indistinguishable (Figure 2C). This immediately poses two questions: First, why does *pag-3* not inhibit *mec-3* expression in ALM? Second, how can *mec-3* antagonize *pag-3* activity in ALM? We will address the first question in this section, showing that *pag-3* operates together with another regulatory system in BDU and for the rest of the paper will address the second question, revealing a novel competition mechanism to explain how *mec-3* antagonizes *pag-3* activity.

Since *pag-3* is required for repression of *mec-3* expression in BDU, but apparently not sufficient to repress *mec-3* expression in ALM (where *pag-3* is also normally expressed), we hypothesized that *pag-3* may operate with other BDU-specific factor(s) to repress *mec-3* expression. This repression occurs on a transcriptional level as the derepression of *mec-3* in *pag-3* is inferred from a transcriptional reporter of the *mec-3* locus (Figure 7A). We considered the possibility that a non-canonical Wnt signal may provide specificity to the *pag-3* mediated repression of *mec-3* expression in BDU. This hypothesis arises from the previously made observation that upon many asymmetric cell divisions along the anterior/posterior axis in the developing embryo, a non-canonical Wnt pathway is activated in the posterior daughter cell (Mizumoto and Sawa, 2007). BDU is the posterior daughter of the embryonic neuroblast division that generates ALM and BDU. To assess whether this non-canonical Wnt signal is indeed active in BDU, we examined two key indicators of the activity of this pathway, the TCF-like protein POP-1 and the β-catenin-like protein SYS-1. In posterior cells in which the Wnt signal is active,

the TCF-like protein POP-1 is exported from the nucleus (Mizumoto and Sawa, 2007), resulting in lower nuclear POP-1 in the posterior nucleus, compared to the anterior nucleus. Moreover, the Wnt signaling system stabilizes the β-catenin -like protein SYS-1 in posterior cells, compared to anterior cells (Mizumoto and Sawa, 2007). We indeed find that after division of the ALM/BDU mother in the embryo, high levels of POP-1 are present in the anterior ALM neuron and low levels in the posterior BDU neuron (Figure 11A). Conversely, we observe low SYS-1 in ALM and high SYS-1 in BDU (Figure 11A).

Figure 11: A non-canonical Wnt signal represses *mec-3* expression in BDU



(A) Expression of *pop-1::gfp* (*qIs74*) and *sys-1::gfp* (*qIs95)* after the ALM/BDU cell division in the embryo. Cells are marked by *unc-86*^{fosmid}*::rfp* (*otEx5851)*.

(B) *mec-3*^{prom}*::gfp (uIs22)* expression is derepressed in BDU in *mom-4(ne1539); lit-1(t1512)* temperature sensitive mutants. Embryos were shifted from the permissive temperature to the restrictive temperature after the birth of the ALM/BDU mother but before the ALM/BDU cell division and analyzed at 2-fold stage. Expression of *mec-3*^{prom}*::gfp* is seen in both ALM and BDU after temperature shift but only in ALM in animals at the permissive temperature. ALM and BDU are marked with arrows.

(C) Schematic summary. The Wnt receptor employed predominantly employed in the "anterior/posterior" coordinate system is MOM-5, the nature (and source) of the ligands is much less explored for many of these signaling events (Mizumoto and Sawa, 2007).

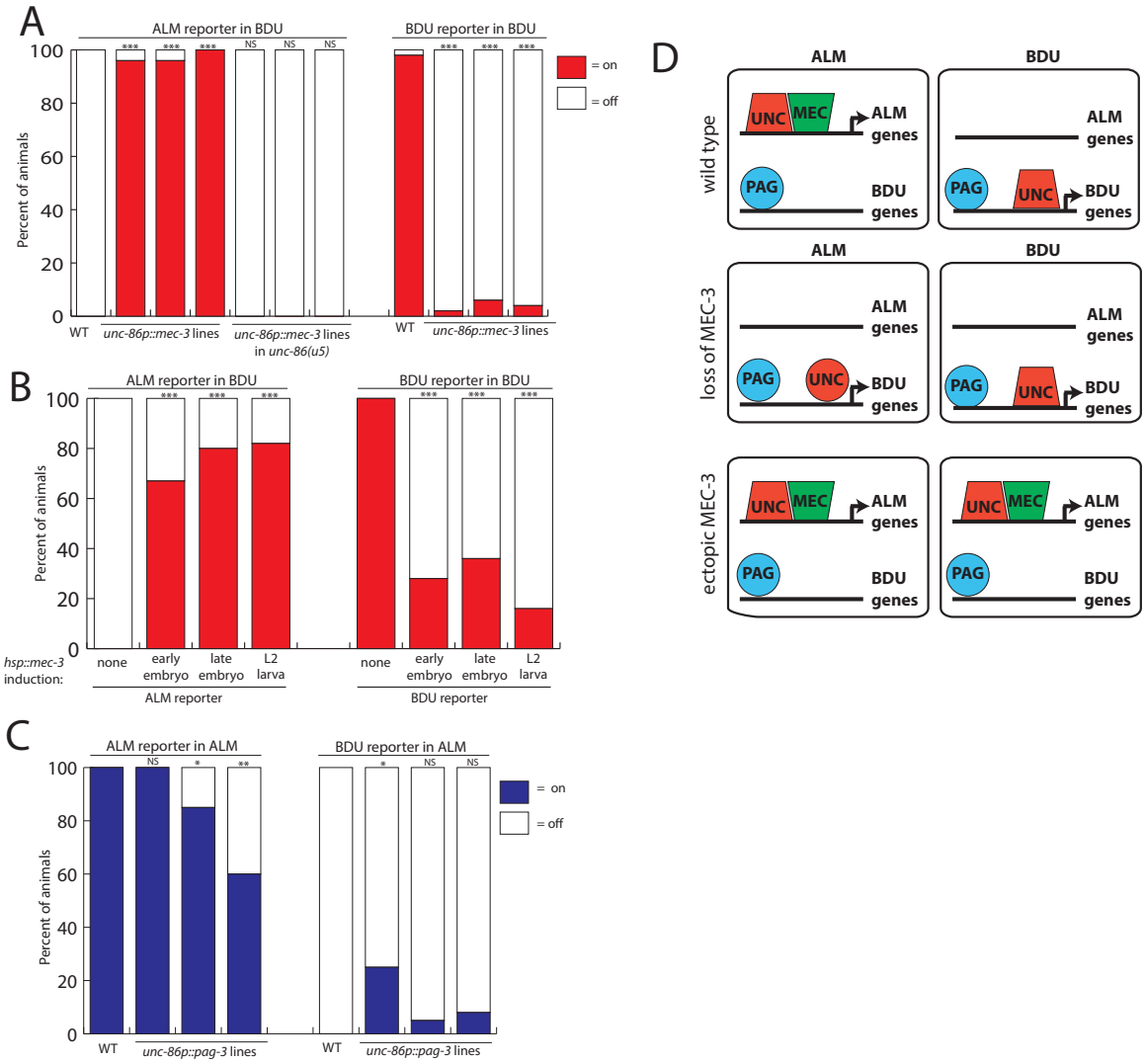
To examine whether the Wnt signaling system in BDU indeed is involved in repressing *mec-3* expression, we altered the activity of the kinases MOM-4 and LIT-1 which are required for the Wnt signaling-dependent export of POP-1 from the posterior sister nucleus (Takeshita and Sawa, 2005). Loss of *mom-4* and *lit-1* function leads to an accumulation of POP-1 in the posterior sister nucleus to a level similar to the one normally observed in the anterior sister. We used the temperature-sensitive double mutant *mom-4(ne1539); lit-1(t1512)* to disrupt the Wnt/ β-catenin asymmetry pathway (Takeshita and Sawa, 2005). *mom-4 (ne1539); lit-1(t1512)* embryos were shifted to the restrictive temperature just before the ALM/BDU division and *mec-3* expression was examined. We find that in temperature-shifted animals, *mec-3* is now derepressed in the posterior BDU neuron (Figure 11B). We conclude that the *pag-3*-dependent, BDU-specific of repression of *mec-3* expression involves a Wnt signaling system (Figure 11C). The Wnt signal may result in the induction of expression of a BDU-specific cofactor with which PAG-3 works together to repress *mec-3* expression; or, alternatively, PAG-3 may cooperate more directly with BDU-enriched SYS-1 to repress *mec-3* expression specifically in BDU.

4.7. MEC-3 outcompetes PAG-3 for UNC-86 access

As mentioned above, the expression of *pag-3* in both BDU and ALM does not only prompt the question how *pag-3* can repress *mec-3* expression in BDU, but also prompts the question of how *mec-3* can antagonize *pag-3* activity in ALM. We first tested whether *mec-3* is not only required to antagonize *pag-3* in ALM, but whether it is also sufficient to antagonize *pag-3* upon ectopic expression in BDU. We found that the

unc-86 promoter-driven *mec-3* expression is able to convert BDU into ALM, as assessed by examination of several terminal identity markers (Figure 12A). The *mec-3*-induced BDU to ALM conversion can be achieved even late (in the second larval stage), long after the two neurons have differentiated in the embryo. We derived this conclusion by inducing *mec-3* in larval stage using the heat-shock promoter (Figure 12B). Overexpression of *pag-3* under control of the heat-shock promoter is, in contrast, not able to convert ALM to BDU (data not shown) and overexpression of *pag-3* under the *unc-86* promoter is also only mildly able to convert ALM into BDU (Figure 12C). We conclude that *mec-3* is a true homeotic gene in the sense that it is not only required to prevent a homeotic transformation, but also sufficient to induce a homeotic transformation (Figure 12D).

Figure 12: Ectopic expression of *mec-3* transforms BDU to ALM



(A,C): Ectopic expression of *mec-3* causes a BDU to ALM transformation, while ectopic expression of *pag-3* does not cause a reciprocal transformation. 5.2kb of the *unc-86* promoter was used to drive expression of *mec-3* (A) or *pag-3* (B). Expression of both an ALM reporter (*mec-17*^{prom}*::rfp;* (*uIs115*) and a BDU reporter (*ceh-14*^{prom}*::gfp,* (*otEx181)* was examined in BDU. Ectopic *mec-3* expression causes *mec-17*^{prom}*::rfp* to turn on in BDU while *ceh-14*^{prom}*::gfp* turns off. The transformation fails to occur in *unc-86(u5)*, which causes a mutation in the MEC-3 binding domain of UNC-86. Animals with altered expression appear similar to those shown in Fig.7.

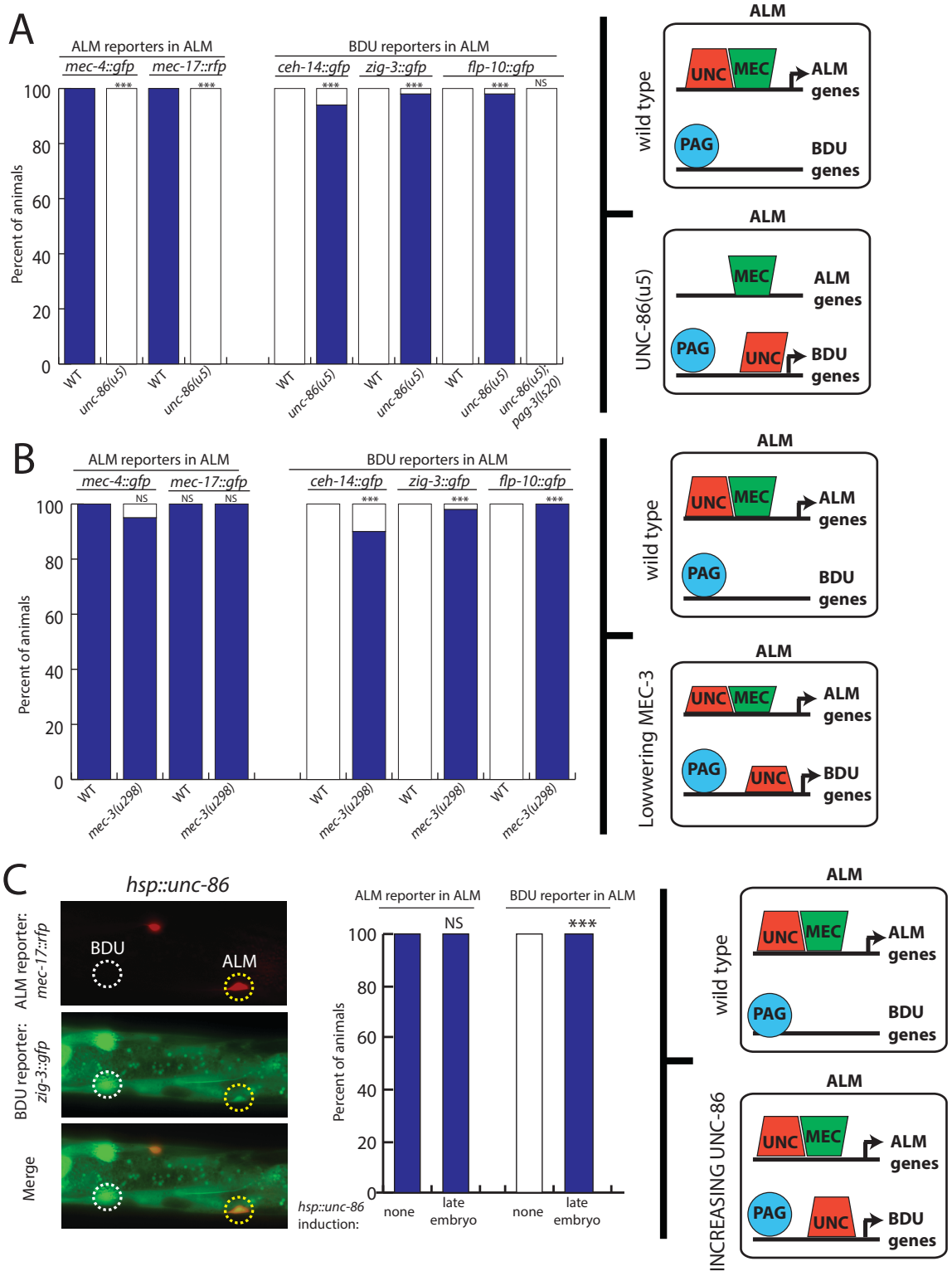
(B) Ectopic expression of *mec-3* causes a BDU to ALM transformation regardless of onset of *mec-3* expression. *mec-3* was expressed under the inducible heat-shock promoter (*hsp-16.2;* (*otEx5852*)) prior to the ALM/BDU cell division (early embryo), during 2- or 3-fold embryonic stages (late embryo), or during L1 or L2 larval stages (larva). Animals were scored 24 hours after heat shock. Reporters used: ALM reporter *mec-17*^{prom}*::rfp*

(*uIs115*) and BDU reporter *ceh-14*^{prom}*::gfp* (*otEx181)*
(D) Schematic summary of the observation that *mec-3* is required in ALM to induce ALM identity and prevent a homeotic transformation to BDU and it is sufficient to convert BDU to ALM identity.
Significance indicates comparison to WT (A,C) or no heat shock (B) using Fischer's Exact Test. ***:$p \leq .0001$, **:$p \leq .001$, *:$p \leq .05$, NS: not significant. For B,C bars show average number of cells. n≥40, with the exception of panel B (n≥20).

Based on the *pag-3* misexpression experiments, we also conclude that that the inability of *pag-3* to induce BDU fate is ALM (i.e., the "recessive" nature of *pag-3* compared to *mec-3*) is not a question of expression levels of *pag-3*. The inability of endogenously expressed *pag-3* to drive BDU identity in ALM could be explained by the presence of an as yet unknown cofactor that is present in BDU, but not ALM. Since ectopic expression of MEC-3 in BDU is able to antagonize UNC-86/PAG-3, we disfavor such a possibility. Instead we considered the possibility that UNC-86 and PAG-3 are sufficient in principle to induce BDU identity in ALM, but are actively prevented from doing so by MEC-3. Two observations lead us to formulate a hypothesis of how MEC-3 may antagonize UNC-86/PAG-3. MEC-3 and UNC-86 directly interact with one another as a heterodimer *in vivo* and *in vitro* and bind to directly adjacent sites on DNA (Röhrig et al., 2000; Xue et al., 1993). In contrast, as we have shown in Figure 6, the UNC-86 and PAG-3 binding sites are spaced by many nucleotides, making heterodimer formation on DNA still possible, but less likely. We therefore hypothesized that MEC-3 may be able to recruit UNC-86 and thereby prevent UNC-86 from cooperating with PAG-3 to induce BDU genes. Two previously described alleles of *unc-86* allowed us to probe this hypothesis. The *unc-86(u5)* and *unc-86(u168)* alleles are missense alleles that were retrieved from screens for touch-insensitive mutants (Chalfie and Au, 1989) and found to affect the physical association of UNC-86 with MEC-3 (Röhrig et al., 2000). The effect of these mutations appear to be selective for the UNC-86/MEC-3 interaction since the ability of UNC-86, which operates with factors other than MEC-3 in other neuron types, to control HSN motor neuron or URX sensory neuron differentiation is unaffected in these mutant animals (Röhrig et al., 2000). We found that in *unc-86(u5)* mutants, the

activation of *unc-86/mec-3*-dependent ALM-specific genes *mec-17* and *mec-4* is indeed disrupted (Figure 13A). The same effect is observed in *u168* mutants (data not shown). However, the BDU-specific genes *ceh-14, flp-10* and *zig-3* are now ectopically activated specifically in ALM (Figure 13A) and this activation depends on *pag-3* (Figure 13A). These results support the hypothesis that if endogenous, wild-type MEC-3 protein is not able to physically interact with UNC-86, the UNC-86 protein will cooperate with PAG-3 to drive BDU fate. These results also indicate that the components to induce BDU identity in ALM are present in ALM but that MEC-3, by specific binding to UNC-86, is able to antagonize PAG-3 and thereby inhibit the execution of BDU fate.

Figure 13: Evidence to support a competition model



(A) Mutations that impair UNC-86/MEC-3 binding cause ALM to BDU transformations. Reporters for ALM are *mec-4^{prom}::gfp (zdIs5)* and *mec-17^{prom}::rfp* (*uIs115*) and BDU reporters are *ceh-14^{prom}::gfp* (*otEx181), zig-3^{prom}::gfp* (*otIs14), flp-10^{prom}::gfp* (*otIs92).* These reporters were crossed into *unc-86(u5)*, causing loss of expression of the ALM reporters and gain of expression of BDU reporters in ALM. Ectopic expression of *flp-10^{prom}::gfp* was eliminated in a *pag-3(ls20)* background.

(B) A *mec-3* hypomorph shows dual expression of ALM and BDU reporters in ALM. ALM and BDU reporters were crossed into *mec-3(u298)*, a mutation with lower expression of *mec-3*. Both ALM and BDU reporters show expression in ALM.

(C) Increased levels of *unc-86* in ALM cause dual expression of ALM and BDU reporters in ALM. *unc-86* was expressed under the inducible heat-shock promoter (*hsp-16.2*) during 3-fold embryonic stage and animals were scored 24 hours later.
Significance indicates comparison to WT (A) or no heat shock (C) using Fischer's Exact Test. ***:$p \leq .0001$, **:$p \leq .001$, *:$p \leq .05$, NS: not significant. In panel A, bars show average number of cells. $n \geq 40$ in panel A,B, n=15 for panel C.

The drawings next to the data panel show a schematic representation of the experimental manipulations. The drawings next to panel A illustrate that UNC-86 is disabled in interacting with MEC-3, but leaving its activity on the BDU target genes intact. The drawing next to panel B illustrate that lowering MEC-3 protein expression (indicated with a smaller circles) now "frees up" part of rate-limiting UNC-86 protein, so that some of the UNC-86 pool can still work together with MEC-3 but some of it can now operate together with PAG-3 on BDU target genes.

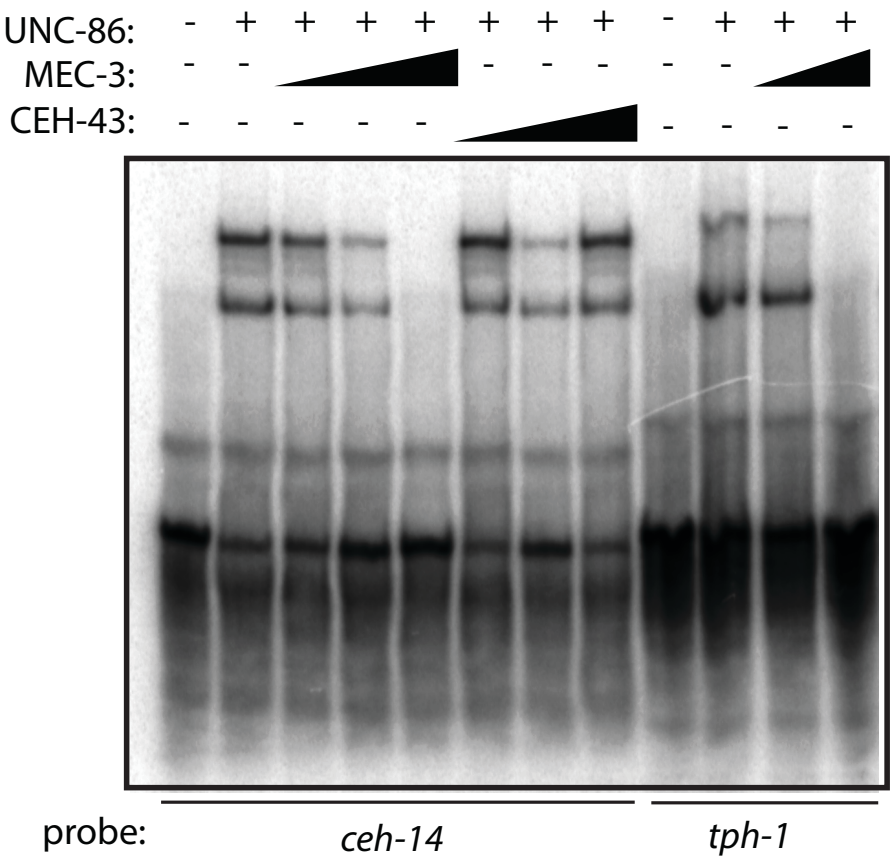
If MEC-3 indeed outcompetes PAG-3 for access to UNC-86, then the BDU to ALM transformation observed upon ectopic expression of *mec-3* in BDU should not occur in an *unc-86(u5)* mutant background in which UNC-86 is not able to interact with MEC-3. We indeed found that ectopic *mec-3* can not induce *mec-17::rfp* expression in BDU in *unc-86(u5)* animals (Figure 12A).

Another prediction can be made based on the model in which MEC-3 competes with PAG-3 to direct all UNC-86 to induce ALM fate, rather than allowing UNC-86 to interact with PAG-3 to generate BDU fates: lowering the level of *mec-3* expression may still provide enough MEC-3 protein to operate together with *unc-86*, but not enough to successfully outcompete PAG-3 binding to UNC-86. In this scenario, ALM markers may still be expressed in ALM, but there may now be ectopic expression of BDU identity markers in ALM as well. We tested this possibility using the *mec-3(u298)* allele, a previously described weak allele with incompletely penetrant mechanosensory defects which contains a transposon insertion upstream of the *mec-3* locus. This allele is thought to lower *mec-3* expression (Way and Chalfie, 1989), a notion we independently confirmed using single molecule fluorescence in situ hybridization (data not shown). We found that in these *mec-3(u298)* mutants, the ALM markers *mec-4* and *mec-17* are still expressed in ALM while the BDU markers *zig-3, ceh-14* and *flp-10* are derepressed in ALM (Figure 13B). This "mixed fate" is different from the null mutant phenotype of *mec-3*, which shows complete loss of ALM markers in ALM, and genetically separates the adoption of ALM identity from the repression of BDU identity. Apparently, different levels of *mec-3* are required for ALM induction and competition with PAG-3 for UNC-86 access.

The competition model predicts that a mixed ALM/BDU fate should be seen not only upon lowering the expression of *mec-3*, but also upon increasing the expression of *unc-86*. Higher levels of *unc-86* expression would provide enough UNC-86 protein to interact with both MEC-3 and PAG-3, thus allowing both sets of identity genes to be expressed simultaneously. We tested this prediction using a heat shock promoter to ubiquitously increase *unc-86* expression beginning at 3-fold stage, after the ALM/BDU division. We found that after induction of *unc*-86, the BDU marker *zig-3* was ectopically expressed in ALM, while expression of the ALM marker *mec-17* was unaffected (Figure 13C).

We used electrophoretic mobility shift assays (EMSAs) to examine the competition model in molecular detail. We find that bacterially produced UNC-86 protein is capable of binding to a 90 bp, double-stranded DNA sequence from the promoter of the BDU-expressed gene *ceh-14*, which contains predicted UNC-86 binding sites. UNC-86 binding to this site can be competed by adding bacterially produced MEC-3 protein, but not by adding equal concentrations of the homeodomain protein CEH-43. We also find that MEC-3 also reduces UNC-86 binding to a 90 bp DNA sequence from the locus of *tph-1* (Figure 14), a gene controlled by UNC-86 and distinct cofactors in other neuron types (Sze et al., 2002; Zhang et al., 2013) (N.Flames and O.H., unpubl. data).

Figure 14: Electrophoretic mobility shift assays reveal impact of MEC-3 on DNA binding by UNC-86



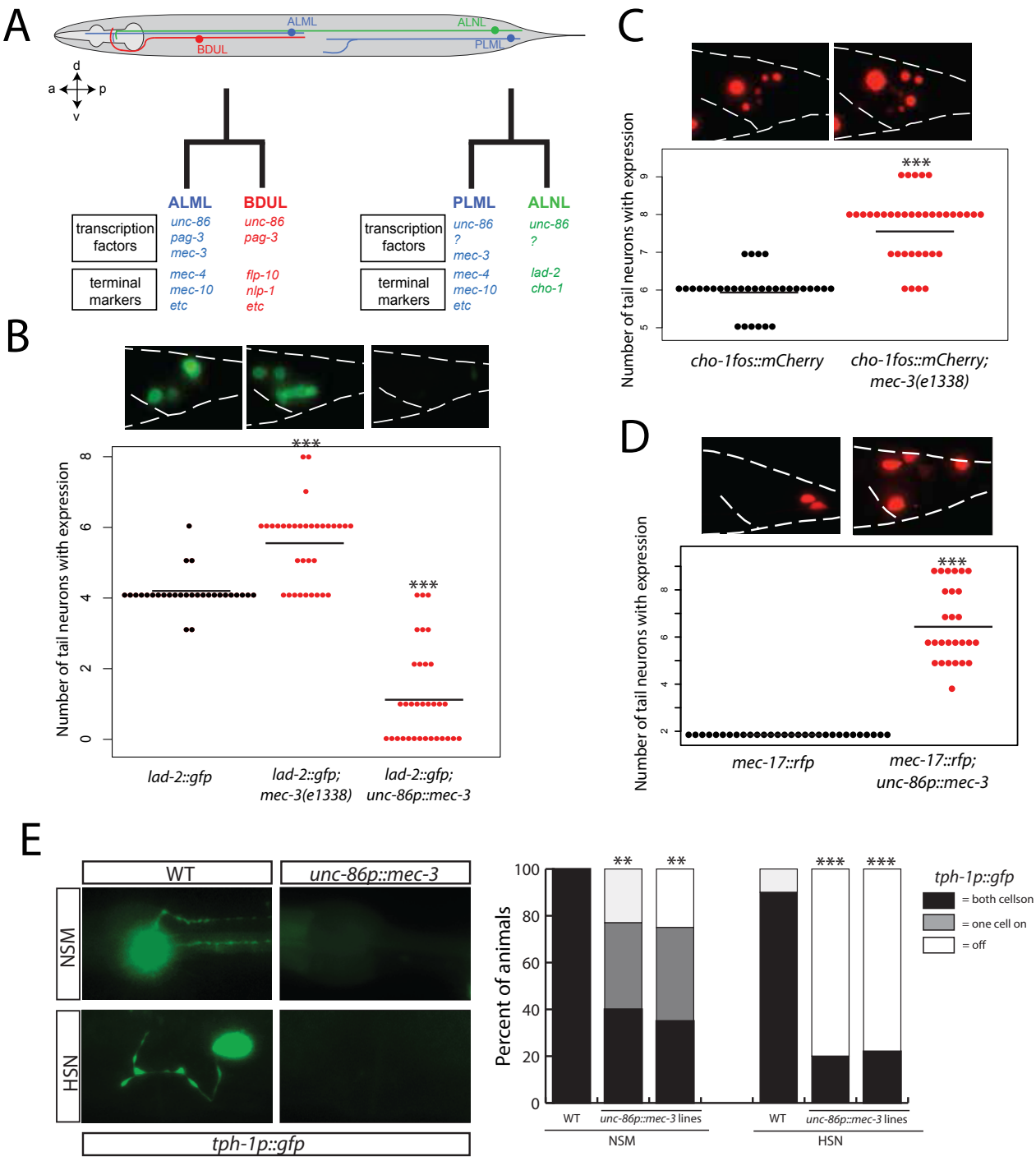
UNC-86 binds to the *mec-3* and *tph-1* promoters; this binding is eliminated with the addition of MEC-3, but not by the addition of the homeodomain protein CEH-43. EMSA was performed with 100 nM UNC-86; 50, 100, and 200 nM (for *ceh-14* probe) or 100 and 200 nM (for *tph-1* probe) MEC-3; and 50, 100, and 200 nM CEH-43. In addition to the probe shown, each lane contained an equal concentration of unlabeled *mec-3* promoter.

Taken together, our data suggests that *unc-86* and *pag-3* drive a "default" BDU state and that this state can, in principle, be induced in both the ALM and BDU neurons. In ALM, however, the presence of *mec-3* diverts from the ground state since *mec-3* can, by direct interaction with UNC-86, prevent the execution of the UNC-86/PAG-3 program and rather induce the ALM differentiation program.

4.8. The competition mechanism operates in other cellular contexts

We tested the generality of the competition mechanism in two different manners. We first considered the PLM/ALN sister neurons in the tail of the animal (Figure 15A). PLM is a light touch receptor neuron that is analogous to ALM in several ways, including its function, overall molecular composition and reliance on the UNC-86/MEC-3 heterodimer for its differentiation (Duggan et al., 1998). Its sister cell is the cholinergic ALN neuron, a neuron that is distinct from the peptidergic BDU neuron, both in terms of overall morphology, synaptic connectivity, molecular profile and neurotransmitter identity (White et al., 1986). However, like the BDU neuron pair, the ALN neuron pair also expresses *unc-86* throughout its lifetime (Finney and Ruvkun, 1990) and requires *unc-86* for its generation (Chalfie et al., 1981). The interaction partner for UNC-86 in ALN is not known.

Figure 15: The competition mechanism operates in other cellular contexts



(A) Schematic drawing of ALM, BDU, ALN, and PLM neurons. Only the left neuron of each pair is shown. Regulatory transcription factors and terminal markers for each neuron are indicated. The question mark in ALNL cannot be *pag-3*, as *pag-3* is not expressed in ALNL (Jia et al., 1997) (data not shown).

(B) *lad-2*^{prom}*::gfp* (*otIs439)* expression is altered by *mec-3* and *unc-86*. Each dot indicates one animal with the given expression of *lad-2*. In *mec-3(e1338)* expression increases by approximately two cells compared to wild type. When *mec-3* is ectopically expressed under the *unc-86* 5.2kb promoter, *lad-2*^{prom}*::gfp* expression decreases variably from zero to four cells.

(C) *cho-1*^{fosmid}*::mCherry* (*otIs544)* expression is altered by *mec-3(e1338)*. In a wild type background, *cho-1*^{fosmid}*::mCherry* is expressed in five to seven cells. In *mec-3(e1338)*, that number increases up to nine cells.

(D) Ectopic *mec-3* causes ALM terminal markers to be ectopically expressed in the tail. Expression of *mec-17*^{prom}*::rfp (uIs115)* increases from two to a variable number between four and nine cells when *mec-3* is ectopically expressed under the *unc-86* 5.2kb promoter. E: Ectopic *mec-3* alters expression of *tph-1*^{prom}*::gfp (zdIs13)*. Expression was examined in HSN and NSM neurons in wild type and with ectopic *mec-3* driven by the *unc-86* 5.2kb promoter. In both neuron types, *tph-1* expression was reduced.

To ask whether in analogy to the ALM/BDU sister neuron pair, *mec-3* also operates in PLM to prevent the a homeotic transformation to ALN fate by competing for UNC-86 access, we tested two predictions: First, in *mec-3* mutants, ectopic expression of ALN markers in PLM should be observed. Using two different marker for ALN identity (choline reuptake transporter *cho-1* and Ig superfamily member *lad-2*), we indeed find this to be the case (Figure 15B,C). Second, ectopic expression of *mec-3* in ALN should convert ALN to PLM identity. Using the *unc-86* promoter to drive *mec-3* in ALN, we indeed find that expression of the ALN marker *lad-2* is abrogated and the PLM marker *mec-17* is ectopically expressed (Figure 15B,D). We conclude that even though the ALN neuron class is very distinct from the BDU neuron class, there are fundamental similarities in the way that their identity is controlled. *unc-86* controls identity of ALN and PLN. In PLM, *mec-3* does not only induce PLM identity but prevents a homeotic transformation to ALN identity, likely by competing with an as yet unknown *unc-86* cofactor expressed in both ALN and PLM. This cofactor (in analogy to *pag-3*) normally drives ALN identity in conjunction with *unc-86* but is prevented by *mec-3* in doing so in PLM.

We further broadened the approach of ectopic *mec-3* expression to assess whether other *unc-86*-dependent cell fate decisions could be disrupted by MEC-3 titrating UNC-86 away from its respective, cell type-specific target genes. We considered the NSM and HSN serotonergic neuron types in which *unc-86* cooperates with a distinct spectrum of cofactors; in NSM, *unc-86* cooperates with the LIM homeobox gene *ttx-3* to drive NSM terminal differentiation (Zhang et al., 2014), while in HSN *unc-86* cooperates with the ETS domain TF *ast-1* and the Zn finger TF *sem-4* (N. Flames and O.H., unpubl. data).

We examined the *unc-86*-dependent expression of *tph-1*, a serotonin biosynthesis pathway gene. We indeed find that ectopic expression of *mec-3* in NSM using the *unc-86* promoter disrupts *tph-1* expression in NSM (Figure 15E). Similarly, *unc-86*-dependent *tph-1* expression in the HSN neurons is also disrupted by ectopic *mec-3* expression (Figure 15E). Lastly, we note that ectopic expression of *mec-3* under control of the *unc-86* promoter not only disrupts the respective differentiation programs of other *unc-86*-expressing neurons, but induces touch marker expression in many of the 57 *unc-86* expressing neurons. For example, ectopic *mec-17* expression can be observed in up to nine additional cells in the tail ganglia of the worm (Figure 15D) which precisely matches the number of *unc-86* expressing neurons in the tail (Finney and Ruvkun, 1990). Taken together, these findings indicate that *mec-3* can operate in very distinct cellular context to "distract" *unc-86* from its normal function and convert neurons into alternative states.

5. Discussion

In the first part of this chapter, we have described gene regulatory routines that define the differentiated state of the BDU neurons. We have shown that the neuropeptidergic identity of BDU, which we found to be critical for its function, constitutes a "subroutine" under control of the *ceh-14* LIM homeobox gene, the *C.elegans* ortholog of vertebrate Lhx3/4. This subroutine is in turn under control of two TFs, the POU homeobox gene *unc-86* (*C.elegans* ortholog of vertebrate Brn3) and the Zn finger TF *pag-3* (*C. elegans* ortholog of vertebrate Gfi), which jointly regulate not only neuropeptidergic identity, but also all other tested molecular identity features of the BDU neurons. The coregulation of distinct identity features by this combination of two transcription factors, likely occurring by direct binding and activation of terminal identity

genes, provides further support for the broad applicability of the concept of neuronal identity control by terminal selectors (Hobert, 2011). As previously observed in several other *C. elegans* neuron classes (Hobert, 2010), terminal selector TFs coregulate many distinct identity features of a specific neuron type. Such coregulation contrasts the alternative, theoretical model of neuronal identity features being controlled in a piece-meal manner by distinct transcription factors.

Previous work had already established that the ALM neurons, the sister neurons of the BDU neurons, are also controlled by two closely cooperating terminal selector-type TFs, *unc-86* and *mec-3* (Chalfie et al., 1981; Duggan et al., 1998; Way and Chalfie, 1989; 1988; Xue et al., 1992; 1993). Notably, the terminal selector combinations for ALM and BDU share a common factor, the *unc-86* POU homeobox gene. Yet the target gene spectrum of UNC-86 is distinct in ALM and BDU and apparently dictated by UNC-86's collaboration with distinct cofactors, the MEC-3 LIM homeodomain protein in ALM and PAG-3 in BDU. Previous work has shown that UNC-86 operates as terminal selector in combination with yet other TFs in completely distinct neuron classes as well, for example, the cholinergic IL2 sensory neurons (Zhang et al., 2014), the serotonergic NSM and HSN neurons (Zhang et al., 2014) or the glutamatergic PVR neurons (Serrano-Saiz et al., 2013).

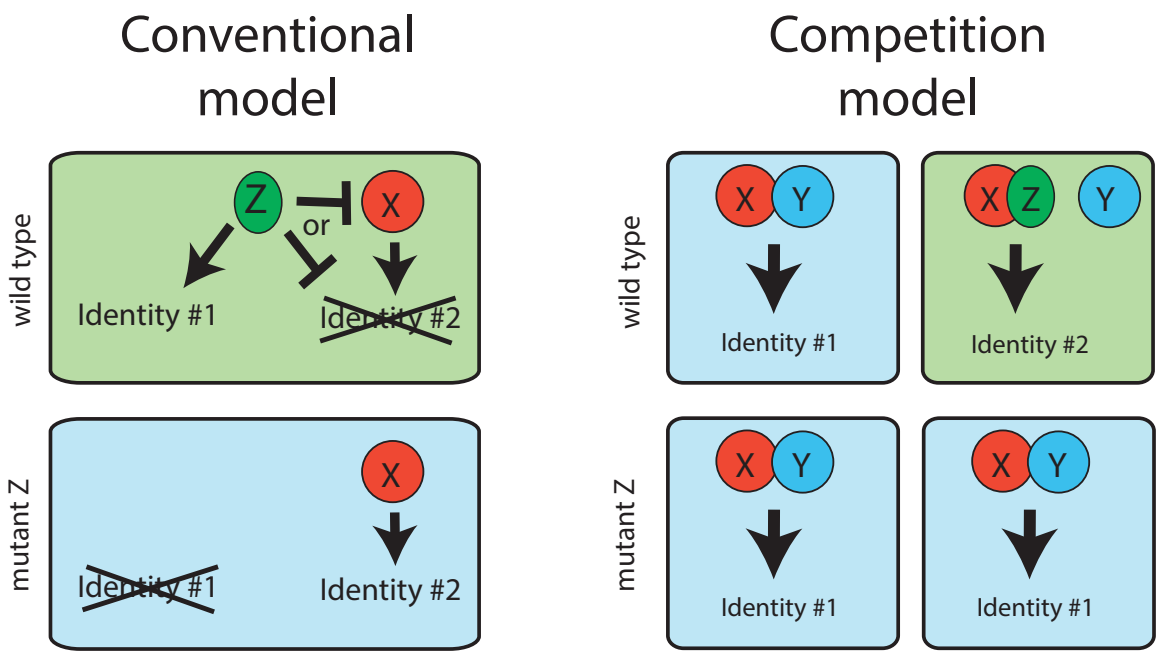
In the second part of this chapter, we have explored the effects of removal of terminal selectors on neuronal identity. Genetic removal of TFs that drive specific neuronal identity programs in either *C. elegans* or other animal species can have remarkably distinct consequences, depending on cellular context; in some cases, neurons will merely remain in an ill-defined, undifferentiated state (e.g. (Altun-Gultekin et al.,

2001; Kratsios et al., 2011; Liu et al., 2010)), in other cases neurons may die (e.g. (Béby et al., 2010)) while in a number of cases, the identity of a neuron switches to the identity of another neuron type (e.g. (Lopes et al., 2012; Sagasti et al., 1999)). Often such identity transformations are just inferred by changes in very select identity features, such as neurotransmitter identity and it is therefore not entirely clear how extensive such transformations are (e.g. (Lopes et al., 2012)). The availability of a host of molecular markers as well as the ability to visualize anatomy in detail allowed us to show that removal of either the BDU terminal selector *pag-3* or the ALM terminal selector *mec-3* results in either complete (*mec-3* mutant) or almost complete (*pag-3* mutant) identity transformations. Corroborating the notion of *mec-3* being a homeotic regulator is our observation that *mec-3* is not only required to prevent a homeotic transformation but also sufficient to promote a homeotic transformation upon ectopic misexpression. The homeotic phenotypes observed in *mec-3* mutants are a testament to the broad impact that a terminal selector (like *mec-3*) has on defining the identity of a neuron type. Homeotic transformation are more conventionally considered in the context of organs or segmented features of an animal but they apparently also extend to transformation of cellular identities, as amply evidenced by the phenotypes of a substantial number of *C. elegans* lineage mutants (Sternberg and Horvitz, 1984).

The conventional interpretation of homeotic identity transformations is that a given transcription factor promotes expression of genes that define one identity, while inhibiting the expression of genes defining an alternative identity (Figure 16; left panels). Indeed in some cases, it has been shown that a TF can have a dual function as an activator and repressor. For example, in the neocortex *Fezf2* directly promotes expression

of *Vglut1*, the key identity determinant of glutamatergic neurons and directly represses expression of *Gad1*, the key identity determinant of GABAergic neurons (Lodato et al., 2014). In the dorsal spinal cord, *Ptf1* directly activates structural identity determinant of GABAergic neurons (e.g. *Gad1*), and directly inhibits TFs that regulate the glutamatergic phenotype (Borromeo et al., 2014). How broadly applicable such dual functionality of a single TF is remains unclear. We have described here evidence that supports a distinct mechanism enabling simultaneous activation of one cellular identity and repression of another. As schematically illustrated in a generalized way in Figure 16 (right panels), MEC-3 induces a differentiation program (ALM identity) by binding to a cooperating TF, UNC-86. Due to limiting amounts of UNC-86 in a cell, this binding makes UNC-86 unable to engage in cooperation with the TF PAG-3. In BDU, UNC-86 is left unperturbed by MEC-3, engages with PAG-3 and induces BDU identity.

Figure 16: Mechanisms to control alternative cell fate choices



Left panels: A commonly observed mechanism for inhibition of alternative cellular identities. Transcription factor Z works to promote one fate while simultaneously inhibiting another fate. This inhibition may either occur through repression of the expression of an inducer of the alternative identity or through direct inhibition of effector genes that define the terminal identity of the respective cell type. Examples exist for both mechanisms (see text for references).

Right panels: An alternative mechanism for inhibition of alternative cellular identities, drawn as a generalized conclusion from our findings. In the case that we have described here "identity #1" is BDU neuron identity, "identity #2" is ALM identity, X is UNC-86, Y is PAG-3 and Z is MEC-3.

The interactions of UNC-86 with MEC-3 and with PAG-3 are likely to be fundamentally distinct. UNC-86 and MEC-3 directly interact with one another and the heterodimer binds to a specific DNA sequence motif with adjacent UNC-86 and MEC-3 binding sites (Duggan et al., 1998; Xue et al., 1993). In contrast, we have shown here that the presumptive UNC-86 and PAG-3 binding sites are physically spaced apart by many dozen nucleotides and their spacing differs in distinct target promoters. Therefore, UNC-86 and PAG-3 may not directly interact in BDU but rather co-conspire to recruit additional factors required for transcriptional activation. Such a scenario applies to many other regulatory elements in which TFs are displayed in a so-called "billboard manner" to recruit transcriptional machinery (Arnosti and Kulkarni, 2005). In this billboard architecture, cooperativity is not a necessity and indeed we find that a removal of multiple UNC-86 and/or PAG-3 sites is required before an effect is observed. By binding directly to UNC-86, MEC-3 is apparently able to not only disrupt this billboard architecture, but also recruit UNC-86 to a distinct set of targets, resulting in a homeotic transformation.

As a note of caution, we have not formally ruled out the possibility that *mec-3* acts by either inducing the expression of a factor in ALM that prevents *unc-86/pag-3* from inducing BDU fate or represses the expression of a BDU-expressed cofactor that *unc-86/pag-3* require to induce BDU fate. We disfavor the postulation of the existence of unknown factors in light of (a) the gene dosage experiments (particularly the *unc-86* dosage), which are much easier to explain with the competition experiments and harder to reconcile with missing cofactors; (b) the gel shift experiments and (c) the observation that *mec-3* can disrupt *unc-86*-mediated differentiation events in completely distinct cellular

contexts that are independent of *pag-3*. For example, *mec-3* acts in PLM to prevent an *unc-86*-dependent homeotic transformation to ALN identity and ectopic *mec-3* expression can disrupt differentiation of the *unc-86*-dependent HSN and NSM neurons. It appears less parsimonious to argue that in all these distinct cellular contexts, some of them entirely artificial (HSN and NSM), *mec-3* is capable of controlling the activity of cell-type specific regulatory cofactors.

The competition mechanism operating in the ALM neurons to prevent a homeotic transformation is contrasted by a distinct mechanism that acts in the BDU neurons to prevent a (partial) homeotic transformation in ALM. This mechanism involves transcriptional repression of a homeotic regulator (*mec-3*), which is mediated by two factors, a Wnt signal and the *pag-3* gene. Based on dual activator/repressor functions of its vertebrate and *Drosophila* orthologs (Jafar-Nejad and Bellen, 2004), we hypothesize that *pag-3* may have a dual role as an activator of the BDU terminal gene battery and a Wnt-dependent repressor of *mec-3* expression. While we have not yet been able to identify *cis*-regulatory element in the *mec-3* locus involved in repressing *mec-3* expression in BDU, a previous study defined *cis*-regulatory elements required for *mec-3* repression in the ALN neurons (Way et al., 1991); this repression is required to prevent ALN to homeotically transform into PLM (this study).

References

Akella, J.S., Wloga, D., Kim, J., Starostina, N.G., Lyons-Abbott, S., Morrissette, N.S., Dougan, S.T., Kipreos, E.T., and Gaertig, J. (2010). MEC-17 is an [α]-tubulin acetyltransferase. Nature *467*, 218–222.

Altun-Gultekin, Z., Andachi, Y., Tsalik, E.L., Pilgrim, D., Kohara, Y., and Hobert, O. (2001). A regulatory cascade of three homeobox genes, ceh-10, ttx-3 and ceh-23, controls cell fate specification of a defined interneuron class in C. elegans. Development *128*, 1951–1969.

Arnosti, D.N., and Kulkarni, M.M. (2005). Transcriptional enhancers: Intelligent enhanceosomes or flexible billboards? Journal of Cellular Biochemistry *94*, 890–898.

Béby, F., Housset, M., Fossat, N., Le Greneur, C., Flamant, F., Godement, P., and Lamonerie, T. (2010). Otx2 Gene Deletion in Adult Mouse Retina Induces Rapid RPE Dystrophy and Slow Photoreceptor Degeneration. PLoS ONE *5*, e11673.

Borromeo, M.D., Meredith, D.M., Castro, D.S., Chang, J.C., Tung, K.C., Guillemot, F., and Johnson, J.E. (2014). A transcription factor network specifying inhibitory versus excitatory neurons in the dorsal spinal cord. Development *141*, 2803–2812.

Cameron, S., Clark, S.G., McDermott, J.B., Aamodt, E., and Horvitz, H.R. (2002). PAG-3, a Zn-finger transcription factor, determines neuroblast fate in C. elegans. Development *129*, 1763–1774.

Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., Frisch, M., Bayerlein, M., and Werner, T. (2005). MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics *21*, 2933–2942.

Cassata, G., Kagoshima, H., Andachi, Y., Kohara, Y., Dürrenberger, M.B., Hall, D.H., and Bürglin, T.R. (2000). The LIM Homeobox Gene ceh-14 Confers Thermosensory Function to the AFD Neurons in Caenorhabditis elegans. Neuron *25*, 587–597.

Chalfie, M., and Au, M. (1989). Genetic control of differentiation of the *Caenorhabditis elegans* touch receptor neurons. Science *243*, 1027–1033.

Chalfie, M., Horvitz, H.R., and Sulston, J.E. (1981). Mutations That Lead to Reiterations in the Cell Lineages of *C. elegans*. Cell *24*, 59–69.

Chelur, D.S., Ernstrom, G.G., Goodman, M.B., Yao, C.A., Chen, L., O' Hagan, R., and Chalfie, M. (2002). The mechanosensory protein MEC-6 is a subunit of the C. elegans touch-cell degenerin channel. Nature *420*, 669–673.

Duggan, A., Ma, C., and Chalfie, M. (1998). Regulation of touch receptor differentiation by the *Caenorhabditis elegans mec-3* and *unc-86* genes. Development *125*, 4107–4119.

Finney, M., and Ruvkun, G. (1990). The *unc-86* Gene Product Couples Cell Lineage and Cell Identity in C. elegans. Cell *63*, 895–905.

Hobert, O. (2011). Regulation of Terminal Differentiation Programs in the Nervous System. Annu. Rev. Cell Dev. Biol. *27*, 681–696.

Jafar-Nejad, H., and Bellen, H.J. (2004). Gfi/Pag-3/Senseless Zinc Finger Proteins: a Unifying Theme? Molecular and Cellular Biology *24*, 8803–8812.

Jia, Y., Xie, G., McDermott, J.B., and Aamodt, E. (1997). The *C. elegans* gene pag-3 is homologous to the zinc finger proto-oncogene gfi-1. Development *124*, 2063–2073.

Jia, Y., Xie, G., and Aamodt, E. (1996). *pag-3*, a *Caenorhabditis elegans* Gene Involved in Touch Neuron Gene Expression and Coordinated Movement. Genetics *142*, 141–147.

Kim, K., and Li, C. (2004). Expression and regulation of an FMRFamide-related neuropeptide gene family in *Caenorhabditis elegans*. J. Comp. Neurol. *475*, 540–550.

Kratsios, P., Stolfi, A., Levine, M., and Hobert, O. (2011). Coordinated regulation of cholinergic motor neuron traits through a conserved terminal selector gene. Nat Neurosci *15*, 205–214.

Lee, R.Y.N., Sawin, E.R., Chalfie, M., Horvitz, H.R., and Avery, L. (1999). EAT-4, a Homolog of a Mammalian Sodium-Dependent Inorganic Phosphate Cotransporter, Is Necessary for Glutamatergic Neurotransmission in *Caenorhabditis elegans*. Journal of Neuroscience *19*, 159–167.

Li, C., and Kim, K. (2010). Neuropeptide Gene Families in *Caenorhabditis elegans*. In Neuropeptide Systems as Targets for Parasite and Pest Control, T.G. Geary, and A.G. Maule, eds. pp. 98–137.

Liu, C., Maejima, T., Wyler, S.C., Casadesus, G., Herlitze, S., and Deneris, E.S. (2010). *Pet-1* is required across different stages of life to regulate serotonergic function. Nat Neurosci *13*, 1190–1198.

Lodato, S., Molyneaux, B.J., Zuccaro, E., Goff, L.A., Chen, H.-H., Yuan, W., Meleski, A., Takahashi, E., Mahony, S., Rinn, J.L., et al. (2014). Gene co-regulation by Fezf2 selects neurotransmitter identity and connectivity of corticospinal neurons. Nat Neurosci *17*, 1046–1054.

Lopes, R., van Wijk, N.V., Neves, G., and Pachnis, V. (2012). Transcription factor LIM homeobox 7 (Lhx7) maintainssubtype identity of cholinergic interneurons in themammalian striatum. Proceedings of the National Academy of Sciences *109*, 3119–3124.

Mizumoto, K., and Sawa, H. (2007). Two βs or not two βs: regulation of asymmetric

division by β-catenin. Trends in Cell Biology *17*, 465–473.

Nathoo, A.N., Moeller, R.A., Westlund, B.A., and Hart, A.C. (2001). Identification of neuropeptide-like protein gene families in Caenorhabditis elegans and other species. Proceedings of the National Academy of Sciences *98*, 14000–14005.

Röhrig, S., Röckelein, I., Donhauser, R., and Baumeister, R. (2000). Protein interaction surface of the POU transcriptionfactor UNC-86 selectively used in touch neurons. The EMBO Journal *19*, 3694–3703.

Sagasti, A., Hobert, O., Troemel, E.R., Ruvkun, G., and Bargmann, C.I. (1999). Alternative olfactory neuron fates are specified by the LIM homeobox gene. Genes & Development *13*, 1794–1806.

Sternberg, P.W., and Horvitz, H.R. (1984). The Genetic Control of Cell Lineage During Nematode Development. Annual Review of Genetics *18*, 489–524.

Sze, J.Y., Zhang, S., Li, J., and Ruvkun, G. (2002). The *C. elegans* POU-domain transcription factor UNC-86 regulates the tph-1 tryptophan hydroxylase gene and neurite outgrowth in specific serotonergic neurons. Development *129*, 3901–3911.

Takeshita, H., and Sawa, H. (2005). Asymmetric cortical and nuclear localizations of WRM-1/β-catenin during asymmetric cell division in C. elegans. Genes & Development *19*, 1743–1748.

Way, J.C., and Chalfie, M. (1989). The mec-3 gene of *Caenorhabditis elegans* requires its own product for maintained expression and is expressed in three neuronal cell types. Genes & Development *3*, 1823–1833.

Way, J.C., Wang, L., Run, J.Q., and Wang, A. (1991). The mec-3 gene contains cis-acting elements mediating positive and negative regulation in cells produced by asymmetric cell division in Caenorhabditis elegans. Genes & Development *5*, 2199–2211.

Way, J.C., and Chalfie, M. (1988). *mec-3*, a homeobox-containing gene that specifies differentiation of the touch receptor neurons in *C. elegans*. Cell *54*, 5–16.

White, J.G., Southgate, E., Thomson, J.N., and Brenner, S. (1986). The Structure of the Nervous System of the Nematode *Caenorhabditis elegans*. Philosophical Transactions of the Royal Society of London B. Biological Sciences *314*, 1–340.

Xue, D., Finney, M., Ruvkun, G., and Chalfie, M. (1992). Regulation of the *mec-3* gene by the *C. elegans* homeoproteins UNC-86 and MEC-3. The EMBO Journal *11*, 4969–4979.

Xue, D., Tu, Y., and Chalfie, M. (1993). Cooperative interactions between the *Caenorhabditis elegans* homeoproteins UNC-86 and MEC-3. Science *261*, 1324–1328.

Zhang, F., Bhattacharya, A., Nelson, J.C., Abe, N., Gordon, P., Lloret-Fernandez, C., Maicas, M., Flames, N., Mann, R.S., Colón-Ramos, D.A., et al. (2014). The LIM and POU homeobox genes *ttx-3* and *unc-86* act as terminal selectors in distinct cholinergic and serotonergic neuron types. Development *141*, 422–435.

Zhang, F., Bhattacharya, A., Nelson, J.C., Abe, N., Gordon, P., Lloret-Fernandez, C., Maicas, M., Flames, N., Mann, R.S., Colon-Ramos, D.A., et al. (2013). The LIM and POU homeobox genes *ttx-3* and *unc-86* act as terminal selectors in distinct cholinergic and serotonergic neuron types. Development *141*, 422–435.

Zweidler-Mckay, P.A., Grimes, H.L., Flubacher, M.M., and Tsichlis, P.N. (1996). Gfi-1 encodes a nuclear zinc finger protein that binds DNA and functions as a transcriptional repressor. Molecular and Cellular Biology *16*, 4024–4034.

CHAPTER 3: CONCLUSION

1. The BDU gene regulatory network

1.1. *unc-86* and *pag*-3 act as terminal selector genes in BDU

Understanding the mechanisms by which different neuron types are formed has been a major goal of the Hobert lab over the past several years. Work by many individual students and postdocs has described the genetic basis for formation of both individual cells and large classes of neurons (Altun-Gultekin et al., 2001; Bertrand and Hobert, 2009; Doitsidou et al., 2013; Etchberger et al., 2007; Flames and Hobert, 2009; Kratsios et al., 2011; Serrano-Saiz et al., 2013; Wenick and Hobert, 2004; Zhang et al., 2014). The sum total of this work presents a picture of complex regulatory decisions that do not all follow one simple pattern but instead exhibit a wide variety of mechanisms. My work fits into this larger picture by examining the regulation of a neuron class that had previously been poorly understood and, more importantly, by placing that regulation in the context of closely related neurons to understand how alternative fates are distinguished.

Four factors have been shown to work together to regulate BDU fate: *unc-86*, *pag-3*, *ceh-14*, and *ahr-1* (this work; Zhang et al., 2013). I have shown in my thesis work that *unc-86* and *pag-3* act as terminal selector genes for BDU: mutants of either gene cause a loss of all BDU identity genes studied, and *cis*-regulatory analysis of the promoters of two BDU-specific genes has revealed functionally important binding sites for both transcription factors. Terminal selectors frequently show auto-regulation (Hobert, 2011), although some terminal selectors, such as *unc-3*, do not appear to auto-regulate (Kratsios et al., 2011). The extent to which *unc-86* regulates its own expression is unclear. A full-length *unc-86* promoter can drive gene expression in all the *unc-86*

neurons even in the background of an *unc-86* mutant, but smaller portions of the promoter are dependent on *unc-86* for expression (Baumeister et al., 1996). It appears, then, that some auto-regulatory elements are present but that other factors may work redundantly with *unc-86* to regulate its expression. *pag-3*, on the other hand, shows negative auto-regulation: Northern analysis of the *pag-3* transcript shows higher levels in a *pag-3* mutant than in wild-type animals (Jia et al., 1997). This analysis was performed on whole-animal extracts, and so it is impossible to learn from it whether *pag-3* regulates itself specifically in BDU neurons.

1.2. Downstream effectors of *unc-86* and *pag-3* regulate BDU subroutines

Downstream of *unc-86* and *pag-3*, two transcription factors regulate specific BDU subroutines. Previous work showed that *ahr-1* controls gap junction formation between BDU and the touch neuron PLM by directly regulating *cam-1*, a receptor tyrosine kinase (Zhang et al., 2013). I found that *ahr-1* expression is significantly reduced in both *unc-86* and *pag-3* mutants, and that *cam-1b* expression is eliminated in the absence of *unc-86*. Because *cam-1* is expressed in ALM as well as BDU, and loss of *pag-3* causes ALM genes to be ectopically expressed in BDU, *cam-1* regulation by *pag-3* could not be assessed.

A second subroutine that we identified in BDU is regulation of neuropeptidergic identity by *ceh-14*. Loss of *ceh-14* causes a partial or total reduction in expression of each of the three neuropeptide genes examined, but it does not affect the majority of the BDU-specific genes we tested. This limited regulatory role for *ceh-14* is interesting, because previous work has shown that *ceh-14* acts as a terminal selector in other contexts (Cassata et al., 2000; Serrano-Saiz et al., 2013). It is possible that *ceh-14* does play a role

in the regulation of all BDU genes, but that its contribution is masked by that of *unc-86* and *pag-3*. Most reporters tested were turned off completely in the absence of *unc-86*, making it impossible to test for enhancement of the phenotype by loss of *ceh-14*. However, a small number of reporters did show only partial loss of expression in the *unc-86* mutant, and it is possible that a *ceh-14;unc-86* double mutant would show a stronger phenotype.

The finding that *ceh-14* specifically regulates neuropeptide identity in BDU is particularly interesting because so little is known about peptide regulation. Most *C. elegans* neurons express at least one neuropeptide gene, and previous work has shown that neuropeptides are among the battery of genes regulated by terminal selectors (Zhang et al., 2014). This work, however, is the first to identify neuropeptides as a specific subroutine regulated by a downstream effector of a terminal selector. Because so little is known about neuropeptide regulation in vertebrates, it is difficult to know whether this regulatory mechanism is found in more complex organisms.

1.3. Combinatorial codes for terminal selectors

One of the most powerful features of a simple model organism such as *C. elegans* is the ability to elucidate the mechanisms by which individual cell fate decisions are made. A remarkable body of work over the past several years has found that the function of even a single transcription factor is vastly altered by the cellular context in which is operates. The regulatory network of BDU exemplifies the complexity of transcription factor interactions. All of the factors known to be expressed in BDU are also found in other *C. elegans* neurons, and they appear to play a different role in each cell type. Part of the explanation for this differential function lies in the combination of transcription

factors that interact in each cell type. The LIM homeobox protein TTX-3, for example, acts as a terminal selector in both AIY and AIA neurons (Zhang et al., 2014). In AIY, TTX-3 interacts with the transcription factor CEH-10, and the heterodimer binds to a distinct DNA motif than TTX-3 alone. In this work, I have described a novel mechanism by which interactions between co-expressed transcription factors can alter the function of those factors. The interaction between UNC-86 and MEC-3 does not merely alter UNC-86 DNA binding, but also prevents UNC-86 from interacting with promoters of heterologous genes. In BDU, the lack of MEC-3 allows UNC-86 to interact with PAG-3 and direct BDU fate.

2. Conservation of the competition mechanism

In chapter 2, I examined a mechanism by which interactions between MEC-3 and UNC-86 alter the function of UNC-86. It is interesting to speculate whether this mechanism exists in other cellular contexts, either in *C. elegans* or in other systems. An obvious starting point for this question is the other *unc-86*-expressing neurons. *unc-86* is found in 57 neurons in adult animals (Baumeister et al., 1996), and it has been shown to be a terminal selector in many of them (Serrano-Saiz et al., 2013; Sze et al., 2002; Zhang et al., 2014). Moreover, *unc-86* interacts with different cofactors in different cell types: in IL2 and URA neurons, it works in conjunction with the ARID transcription factor *cfi-1*, while in NSM neurons, it works with *ttx-3* (Zhang et al., 2014). It is not clear in all of these cases whether UNC-86 is physically interacting with the different cofactors and if a sequestration mechanism is occurring in these cells. Given the wide range of genes that can be bound by UNC-86, there must be some mechanism for inhibiting alternative binding in each cell type. Certainly, sequestration is not the only possible means of preventing *unc-86* from inducing the incorrect cell fate. Chromatin modifications or the

repressive action of other transcription factors could also limit *unc-86* function in each neuron type. More work needs to be done to understand how UNC-86 differentially regulates each of neuron.

Several studies of *Drosophila* homeotic transformation have indicated a competition mechanism. This mechanism has particularly been noted with regard to the "posterior dominance" phenomenon, in which ectopic expression of posteriorly-expressed homeotic genes into anterior regions causes a transformation but the reverse does not. For example, the thoracic transcription factor *Antennapedia* (*Antp*) causes a loss of eye fate and an antenna-to-leg transformation when expressed in the anterior imaginal discs. These two phenotypes can be uncoupled, and mutant analysis showed that the transformation requires *Antp* DNA binding capabilities, while the eye loss requires physical interaction between *Antp* and the eye selector gene *Eyeless* (Plaza et al., 2008). Similarly, the posterior Hox protein Abdominal-A (AbdA) inhibits the activity of *Sex combs reduced* by competing for DNA binding, a competition that requires interaction with the Hox cofactor Extradenticle (Noro et al., 2011). These interactions, much like the MEC-3/UNC-86 heterodimer, alter cell fate. A key distinction is that the *Drosophila* Hox protein interactions appear to occur only upon ectopic expression of a posterior factor, whereas MEC-3 and UNC-86 interact endogenously in ALM. Thus while transcription factor binding may cause homeotic transformations in the case of ectopic expression of *Drosophila* Hox genes, it prevents homeotic transformation in ALM neurons.

Although most *Drosophila* Hox genes are expressed in separate segments of the animal, some overlap does occur. For example, *Ultrabithorax* (*Ubx*) and *AbdA* are

coexpressed in parts of the epiderm. While they do not fall into a clear dominance hierarchy, overexpression experiments have shown that the more-posterior *AbdA* has a somewhat more dominant role (Duboule and Morata, 1994). Given that physical interaction between Hox proteins or their cofactors is necessary for fate changes upon misexpression of posterior genes, it is possible that a similar mechanism is at work in the case of the coexpressed genes. More work needs to be done to determine whether that is indeed the case.

Another prominent example of the function of a transcription factor being altered by interactions with its partners comes from the yeast *Saccharomyces cerevisiae*. This unicellular organism can form three different cell types: the haploid a and α cells and the diploid a/α (Lodish et al., 2000). Specification of these three cell types is mediated by three cell type-specific transcription factors called a1, α1, and α2, as well as one ubiquitously expressed factor, MCM1. It is this last factor whose variable selectivity allows transcription of a different set of cell type-specific genes. MCM1 binds well to the upstream regulatory sequences (URSs) of a-specific genes, but it binds poorly to the URSs of α-specific genes. In α-type cells, however, MCM1 interacts with α1 to bind to the promoters of α-type genes with high affinity. In these changing affinities, the interactions between MCM1 and α1 resemble those of UNC-86 and MEC-3. The parallels are not exact, however. While MEC-3 physically sequesters UNC-86 away from the BDU promoters, it does not appear that any other factors are simultaneously inhibiting expression of BDU-specific genes. In α cells, however, the factor α2 interacts with MCM1 on the promoters of a-specific genes and causes a repression of transcription.

This complex similarly functions to repress a-specific gene expression in a/α diploid cells, while the factor a1 works in diploid cells to repress expression of α1.

While the competition mechanism we have shown here has never been definitively shown to occur in vertebrates, several studies have presented intriguing possibilities. In the chick dorsal horn, *Lbx1* and *Tlx3* act as opposing switches for GABAergic and glutamatergic fate, respectively (Cheng et al., 2005). Intriguingly, *Lbx1* is expressed in both neuron types. Loss of *Tlx3* causes a transformation from glutamatergic to GABAergic fate, but that switch is dependent on *Lbx1* expression. Moreover, higher levels of *Lbx1* expression can cause presumptive glutamatergic cells to adopt a GABAergic fate, even in the presence of *Tbx3*. Taken together, these results suggest a sequestration mechanism in which *Tlx3* directly or indirectly antagonizes *Lbx1*. To fully establish this interaction as equivalent to the physical sequestration of UNC-86 by MEC-3, it would necessary to first find evidence for direct interaction between *Lbx1* and *Tlx3* (or one of its downstream targets) and then show that the interaction alters *Lbx1* DNA binding capabilities.

3. A "default" BDU state

One of the enduring mysteries of this research was the question of why *pag-3* is expressed in ALM and the other mechanosensory neurons, when it appears to only function in ALM in the absence of *mec-3*. It is possible that *pag-3* does in fact play a role in ALM development. Gentle touch response in the *pag-3* mutant cannot be tested by behavioral assays because *pag-3* is also expressed in the ventral nerve cord, and the mutant has backwards movement defects. ALM signaling could be more directly tested using calcium imaging. However, given that loss of *pag-3* does not cause a decrease in

expression of any members of the ALM gene battery thus far tested, it seems unlikely that the mutant would show any loss of mechanosensory response.

An alternative explanation for the expression of *pag-3* in ALM comes from an examination of the evolutionary implications of homeosis. Homeotic transformation has long been seen as a driver of rapid evolutionary change, as mutations to regulatory factors can cause wide-scale alterations in developing cells (Sternberg and Horvitz, 1984). Alternatively, changes in either the expression of homeotic factors or the complex *cis*-regulatory architecture of their target genes allow small-scale tissue alterations that would have a lower chance of having adverse pleiotropies (Akam, 1998; Stern, 2000). This idea can be applied to ALM and BDU in the context of the "BDU default" state. Both ALM and BDU neurons are competent to differentiate into a BDU-type cell, with all the factors necessary (*unc-86* and *pag-3*) already expressed in ALM. Thus BDU is the default neuronal fate for both cells, but the addition of *mec-3* in ALM causes the neuron to instead adopt ALM fate. It is possible that an ancestral nematode formed two BDU cells instead of one BDU and one ALM. Changes to either *unc-86* or the promoter of *mec-3* caused *mec-3* to be expressed in ALM and the other mechanosensory neurons, thus altering them and forming a new class of neurons. This is of course only speculation, but it is convincing in light of the interaction between MEC-3 and PAG-3 in ALM.

4. Neuropeptide-mediated function of BDU

While previous work had shown that BDU plays a role in the harsh touch circuitry of *C. elegans* (Li et al., 2011), the genetic basis by which BDU mediated touch response was unclear. BDU does not directly synapse onto any neurons that direct backwards movement (White et al., 1986), which raised the possibility that it signals through non-synaptic release of neuropeptides. In chapter 2, I presented evidence that removal of the

neuropeptide *nlp-1* causes defects in harsh touch, and that those defects can be rescued by BDU-specific expression of *nlp-1*. Two important questions remain. The first is whether BDU is acting as an interneuron or if it is a sensory neuron that directly senses harsh touch. This latter possibility is intriguing, as BDU has a long projection that runs along almost the entire anterior half of the animal. The sensory or interneuron nature of BDU could be determined either by identifying some other harsh touch sensory neuron that signals to BDU, indicating it acts as an interneuron, or by finding mechanoreceptors in BDU that are required for harsh touch response.

The second question that still remains is which neuron BDU is signaling to. This question is difficult to answer for a number of reasons. First, the possibility of long-range neuroptide signaling means that proximity and synaptic connections cannot necessarily be used to determine which neurons are downstream of BDU in the harsh touch circuitry. A list of candidates could be made from the neurons whose ablation causes defects in harsh touch response (Li et al., 2011). The best way to test those candidates would be to examine mutants of neuropeptide receptors that are normally expressed in each neuron. This brings up a second difficulty: while the neuropeptide genes are very well characterized, neuropeptide receptors have been less studied. Bioinformatic analysis has suggested 153 putative neuropeptide receptors (Hobert, 2013), although more may be present. The expression pattern of a majority of those potential receptors has not been studied. Moreover, the neuropeptide-receptor pairs have not been determined for most neuropeptides. Overall, a greater understanding of *C. elegans* neuropeptide signaling is needed in order to fully understand BDU-mediated harsh touch response.

While at least one neuropeptide in BDU is involved in harsh touch circuitry, other neuropeptides, such as *nlp-15*, do not appear to have an effect on touch response. The *nlp-15* mutant did not have any obvious phenotype to distinguish it from wild type, but more careful analysis may indicate some function for the peptide. Alternatively, *nlp-15* may act redundantly with other signaling molecules, either in response to harsh touch or to mediate some other function of BDU.

5. Remaining questions and future directions

Do ALM and BDU adopt reciprocal functions upon homeotic transformation?

While BDU exhibits only a partial homeotic transformation in terms of morphology changes, all ALM genes tested are expressed in BDU upon loss of *pag-3*. Those genes code for functionally important parts of the mechanosensory neurons, including the DEG/ENaC channel (MEC-4), the specialized microtubules (MEC-7), and glutamate transport (EAT-4). Additionally, translational fusions of MEC-4 indicated a punctate localization, suggesting that the sodium channel is properly assembled and transported along the axon. Because of these results, it appears likely that the BDU-to-ALM transformation is functionally complete; that is, BDU is able to sense and transmit signals in response to gentle touch. The best way to test this would be by using a calcium indicator such as GCaMP. It is important to note, however, that even if BDU acts like a mechanosensory neuron in *pag-3* mutants, it is likely that it could not function in place of ALM to allow anterior gentle touch response. Because BDU morphology is mostly unchanged in *pag-3*, the transformed BDU is physically unable to make the same synaptic connections as ALM. As such, it may be acting like a mechanosensory neuron but failing to connect to the rest of the touch response circuitry. As *pag-3* causes

backwards movement defects as well as BDU transformations, this idea is hard to test by behavioral assays.

The reciprocal question is whether ALM behaves like BDU in addition to looking like BDU in *mec-3* mutants. In this case, the homeotic transformation is complete: ALM not only expresses the BDU gene battery, but also is positioned more ventrally and medially such that its axon follows the same fascicle as BDU. The transformed ALM may thus be correctly placed to recapitulate BDU function, and could be expected to compensate for loss of BDU. As the harsh touch defects associated with loss of BDU are relatively subtle and other functions of BDU have not been well-characterized, these experiments would need to be done with a great deal of precision.

What relative levels of mec-3 and unc-86 expression are required in ALM to cause sequestration?

The work I presented in this thesis showed that MEC-3 physically interacts with UNC-86 in ALM to sequester it away from BDU gene promoters. What is not yet clear from my experiments is the stoichiometry of this interaction. Two questions arise: how efficient is MEC-3 at sequestering UNC-86, and how much free UNC-86 is sufficient to induce a dual ALM/BDU fate? As I discuss in Appendix A, several overexpression experiments with *unc-86* failed to cause ectopic expression of BDU genes. Injection of neither the *unc-86* fosmid nor amplified *unc-86* genomic locus were sufficient to turn on BDU genes in ALM. Instead, the dual fate was only seen after using the strong heat shock promoter to drive *unc-86* expression. These results suggest that a large change in the ratio of MEC-3:UNC-86 is necessary to create a dual fate in ALM. That conclusion

comes with a strong caveat because it is based on negative results. However, the idea is borne out by the *mec-3* null heterozygote, which similarly did not show any ectopic expression of BDU genes.

The simplest explanation for the robustness of *mec-3*-mediated inhibition of BDU genes is that ALM normally contains a large excess of MEC-3. To test this explanation, it would be necessary to more firmly establish the levels of *unc-86* and *mec-3* in ALM. Transcript levels in particular cells can be precisely measured using smFISH, although that does not account for half-life of the proteins and so cannot necessarily give an accurate estimate of the total protein amount.

What is the mechanism by which mec-3 is restricted from BDU?

While *mec-3* is normally expressed in ALM, loss of either *pag-3* or Wnt signaling causes ectopic expression in BDU. Thus far, I have been unable to determine the mechanism for repression of *mec-3*. Both previous work (Way et al., 1991) and my own experiments (Appendix A) have found that removal of particular sections of the *mec-3* promoter results in ectopic expression in sister cells. However, none of those deletions or mutations, including removal of putative *pag-3* binding sites, resulted in BDU expression. Despite this failure to identify a *cis*-regulatory element required for repression, it seems likely that either PAG-3 or one of its downstream targets is directly binding to the *mec-3* promoter to inhibit expression in BDU. Given the dual activator/repressor role of the *pag-3* homologs, it would not be surprising if *pag-3* is the direct repressor of *mec-3*. In that case, we hypothesize that there may be some Wnt-dependent cofactor that interacts with PAG-3 to effect repression.

References

Akam, M. (1998). Hox genes, homeosis and the evolution of segment identity: no need for hopeless monsters. Int. J. Dev. Biol. *42*, 445–451.

Altun-Gultekin, Z., Andachi, Y., Tsalik, E.L., Pilgrim, D., Kohara, Y., and Hobert, O. (2001). A regulatory cascade of three homeobox genes, ceh-10, ttx-3 and ceh-23, controls cell fate specification of a defined interneuron class in C. elegans. Development *128*, 1951–1969.

Baumeister, R., Liu, Y., and Ruvkun, G. (1996). Lineage-specific regulators couple cell lineage asymmetry to the transcription of the *Caenorhabditis elegans* POU gene unc-86 during neurogenesis. Genes & Development *10*, 1395–1410.

Bertrand, V., and Hobert, O. (2009). Linking Asymmetric Cell Division to the Terminal Differentiation Program of Postmitotic Neurons in *C. elegans*. Developmental Cell *16*, 563–575.

Cassata, G., Kagoshima, H., Andachi, Y., Kohara, Y., Dürrenberger, M.B., Hall, D.H., and Bürglin, T.R. (2000). The LIM Homeobox Gene ceh-14 Confers Thermosensory Function to the AFD Neurons in Caenorhabditis elegans. Neuron *25*, 587–597.

Cheng, L., Samad, O.A., Xu, Y., Mizuguchi, R., Luo, P., Shirasawa, S., Goulding, M., and Ma, Q. (2005). Lbx1 and Tlx3 are opposing switches in determining GABAergic versus glutamatergic transmitter phenotypes. Nat Neurosci *8*, 1510–1515.

Doitsidou, M., Flames, N., Topalidou, I., Abe, N., Felton, T., Remesal, L., Popovitchenko, T., Mann, R.S., Chalfie, M., and Hobert, O. (2013). A combinatorial regulatory signature controls terminal differentiation of the dopaminergic nervous system in *C. elegans*. Genes & Development *27*, 1391–1405.

Duboule, D., and Morata, G. (1994). Colinearity and functional hierarchy among genes of the homeotic complexes. Trends in Genetics *10*, 358–364.

Etchberger, J.F., Lorch, A., Sleumer, M.C., Zapf, R., Jones, S.J., Marra, M.A., Holt, R.A., Moerman, D.G., and Hobert, O. (2007). The molecular signature and cis-regulatory architecture of a *C. elegans* gustatory neuron. Genes & Development *21*, 1653–1674.

Flames, N., and Hobert, O. (2009). Gene regulatory logic of dopamine neuron differentiation. Nature *458*, 885–889.

Hobert, O. (2011). Regulation of Terminal Differentiation Programs in the Nervous System. Annu. Rev. Cell Dev. Biol. *27*, 681–696.

Hobert, O. (2013). The neuronal genome of *Caenorhabditis elegans*. WormBook.

Jia, Y., Xie, G., McDermott, J.B., and Aamodt, E. (1997). The *C. elegans* gene pag-3 is homologous to the zinc finger proto-oncogene gfi-1. Development *124*, 2063–2073.

Kratsios, P., Stolfi, A., Levine, M., and Hobert, O. (2011). Coordinated regulation of cholinergic motor neuron traits through a conserved terminal selector gene. Nat Neurosci *15*, 205–214.

Li, W., Kang, L., Piggott, B.J., Feng, Z., and Xu, X.Z.S. (2011). The neural circuits and sensory channels mediating harsh touch sensation in *Caenorhabditis elegans*. Nat Comms *2*, 315.

Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, Pl, Baltimore, D., and Darnell, J. (2000). Cell-type specification and mating-type conversion in yeast. Molecular Biology of the Cell, 4th edition.

Noro, B., Lelli, K., Sun, L., and Mann, R.S. (2011). Competition for cofactor-dependent DNA binding underlies Hox phenotypic suppression. Genes & Development *25*, 2327–2332.

Plaza, S., Prince, F., Adachi, Y., Punzo, C., Cribbs, D.L., and Gehring, W.J. (2008). Cross-regulatory protein–protein interactions between Hox and Pax transcription factors. Proceedings of the National Academy of Sciences *105*, 13439–13444.

Serrano-Saiz, E., Poole, R.J., Felton, T., Zhang, F., La Cruz, De, E.D., and Hobert, O. (2013). Modular Control of Glutamatergic Neuronal Identity in *C. elegans* by Distinct Homeodomain Proteins. Cell *155*, 659–673.

Stern, D.L. (2000). Perspective: evolutionary developmental biology and the problem of variation. Evolution *54*, 1079–1091.

Sternberg, P.W., and Horvitz, H.R. (1984). The Genetic Control of Cell Lineage During Nematode Development. Annual Review of Genetics *18*, 489–524.

Sze, J.Y., Zhang, S., Li, J., and Ruvkun, G. (2002). The *C. elegans* POU-domain transcription factor UNC-86 regulates the tph-1 tryptophan hydroxylase gene and neurite outgrowth in specific serotonergic neurons. Development *129*, 3901–3911.

Way, J.C., Wang, L., Run, J.Q., and Wang, A. (1991). The mec-3 gene contains cis-acting elements mediating positive and negative regulation in cells produced by asymmetric cell division in Caenorhabditis elegans. Genes & Development *5*, 2199–2211.

Wenick, A.S., and Hobert, O. (2004). Genomic cis-Regulatory Architecture and trans-Acting Regulators of a Single Interneuron-Specific Gene Battery in *C. elegans*. Developmental Cell *6*, 757–770.

White, J.G., Southgate, E., Thomson, J.N., and Brenner, S. (1986). The Structure of the Nervous System of the Nematode *Caenorhabditis elegans*. Philosophical Transactions of the Royal Society of London B. Biological Sciences *314*, 1–340.

Zhang, F., Bhattacharya, A., Nelson, J.C., Abe, N., Gordon, P., Lloret-Fernandez, C.,

Maicas, M., Flames, N., Mann, R.S., Colón-Ramos, D.A., et al. (2014). The LIM and POU homeobox genes *ttx-3* and *unc-86* act as terminal selectors in distinct cholinergic and serotonergic neuron types. Development *141*, 422–435.

Zhang, J., Li, X., Jevince, A.R., Guan, L., Wang, J., Hall, D.H., Huang, X., and Ding, M. (2013). Neuronal Target Identification Requires AHA-1-Mediated Fine-Tuning of Wnt Signaling in C. elegans. PLoS Genet *9*, e1003618.

APPENDIX A: Table of mutant phenotypes

Gene	Mutant allele	Phenotype
unc-86	*u168*	Loss of expression of ALM- and BDU-specific genes
	u5	Ectopic expression of BDU genes in ALM
mec-3	*e1338*	Transformation of ALM to BDU fate
	u298	Dual ALM/BDU fate adopted in ALM
pag-3	*ls20*	Transformation of BDU to ALM fate
ceh-14	*ch3*	Loss of neuropeptide expression in BDU

APPENDIX B: Regulation of *unc-86*-expressing neurons

Regulation of URA and AIZ neuron fate by unc-86

While my research focused on ALM and BDU fate determination, my project began with a broader analysis of *unc-86* expressing neurons to determine if *unc-86* regulated the terminal gene battery of each neuron type. I began by looking at a few markers in each of three different neuron classes: IL2, URA, and AIZ. My work on IL2 was incorporated into a larger manuscript by Feifan Zhang, and is shown in Appendix B.

AIZ neurons are a left/right bilateral pair with cell bodies in the posterior pharynx and projections into the nerve ring (White et al., 1986). *unc-86* is expressed in AIZ (Baumeister et al., 1996), but its role in the neuron is unclear. I examined reporters for two genes expressed in AIZ: *odr-2* and *ser-2*. After crossing each strain into *unc-86(n846)*, I found that both reporters were no longer expressed in AIZ. While these data are preliminary, it appears that *unc-86* does play a role in regulation of AIZ neurons.

URA neurons are four cholinergic neurons found in the head of the animal (White et al., 1986). In URA, the ARID transcription factor *cfi-1* inhibits *unc-86*-mediated ectopic expression of CEM neuron genes (Shaham and Bargmann, 2002). It was unclear, however, whether *unc-86* directly regulates URA genes themselves. I examined three terminal markers of URA neurons: the neuropeptide genes *flp-21* and *flp-22*, and the glutamate receptor *glr-4*. I found that expression of *flp-21* and *flp-22* were unaffected in *unc-86(n846)*, but that *glr-4* expression was drastically reduced. Later work by Feifan Zhang showed that cholinergic gene battery expression was also reduced in *unc-86* mutants (Zhang et al., 2014).

A forward genetic screen for repressors of mec-3

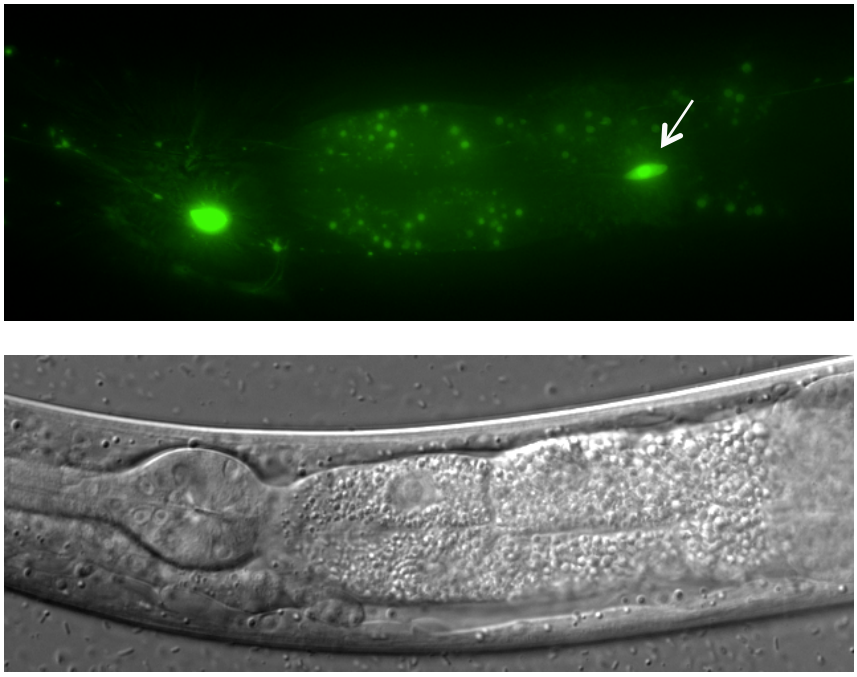
In order to find other factors besides *pag-3* that could be repressing *mec-3* expression in BDU, I undertook a forward genetic screen with the help of a high-school summer student, Angela Kim. We used as a starting strain *uIs22*, a *mec-3::gfp* reporter that does show ectopic expression in BDU in *pag-3* mutants. After mutagenizing with ethyl methane sulfonate (EMS), we performed a semi-clonal screen looking for changes in expression of *mec-3::gfp*. After screening through 840 genomes, we identified six mutants that fell into three classes (Table 1). One mutant, *ot698*, showed ectopic expression in BDU (Figure 1A), while five mutants, *ot699*-*ot703*, fell into one of two different classes: they either had ectopic expression in one anterior neuron that was not BDU (*ot699* and *ot702*) (Figure 2A), or they had ectopic expression in one or two tail cells (*ot700*, *ot701*, and *ot703*) (Figure 3A). The mutant of greatest interest to us was *ot698*. In this mutant, *uIs22* is expressed in BDU beginning at L2-stage (Figure 1B), although it shows expression in ALM beginning at embryonic stages. This delayed onset is similar to that seen in *pag-3*(*ls20*). In addition, the mutant showed the same reverse kinker Unc phenotype as *pag-3*(*ls20*), suggesting that it could be a *pag-3* allele. Complementation testing revealed that to be the case.

Table 1: Classes of mutants found in a screen for repressors of *mec-3*

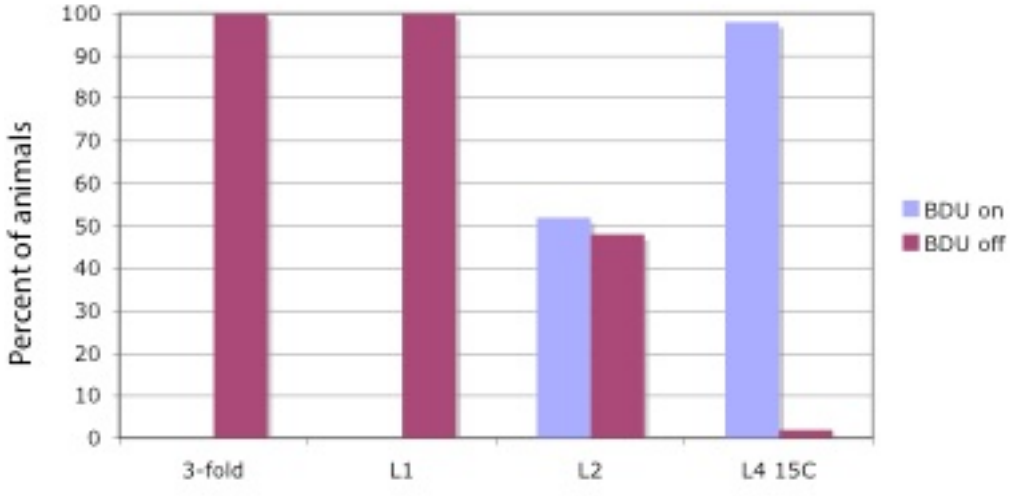
Class	Phenotype	Genotype
A	Expression in BDU	*ot698*
B	Expression in one anterior neuron	*ot699*
		ot702
C	Expression in one or two tail cells	*ot700*
		ot701

Figure 1: *ot698* shows ectopic expression of *mec-3* in BDU neurons

(A)



(B)

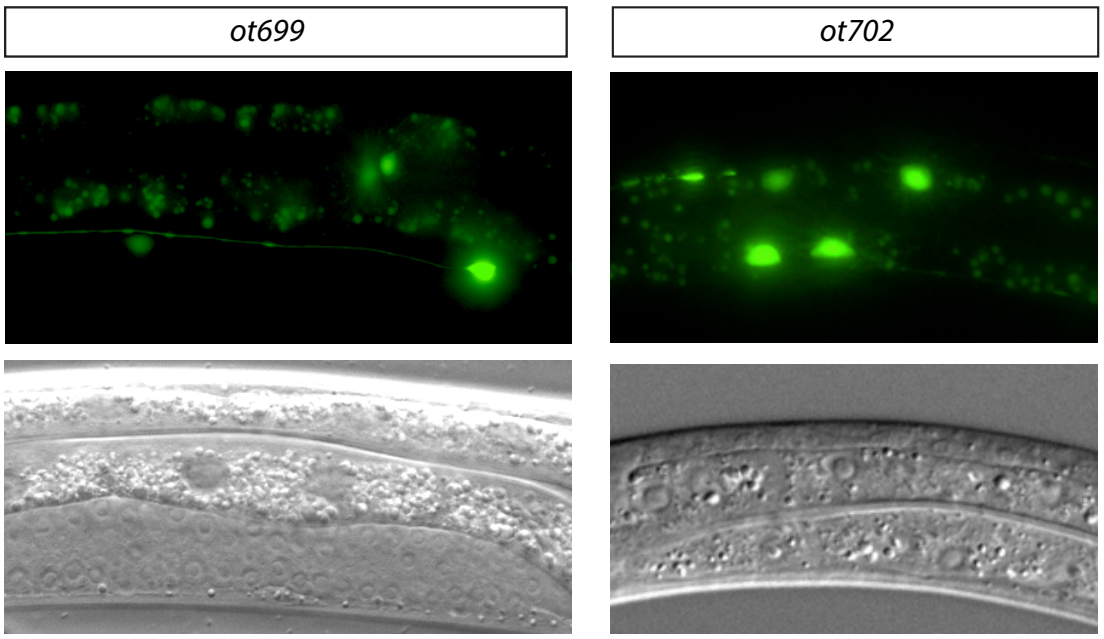


(A) *mec-3::gfp* expression in BDU. BDU is indicated by the white arrow.

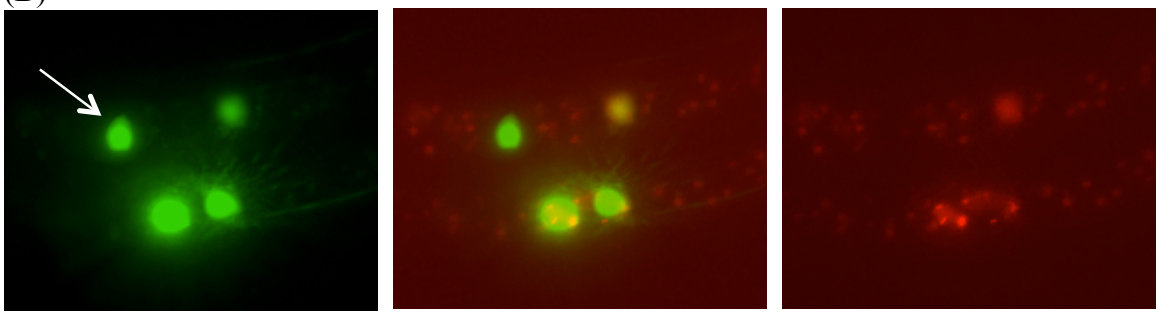
(B) Onset of *mec-3::gfp* expression in BDU. For each time point, $n \geq 40$.

Figure 2: Ectopic expression of *mec-3::gfp* in Class B mutants

(A)



(B)

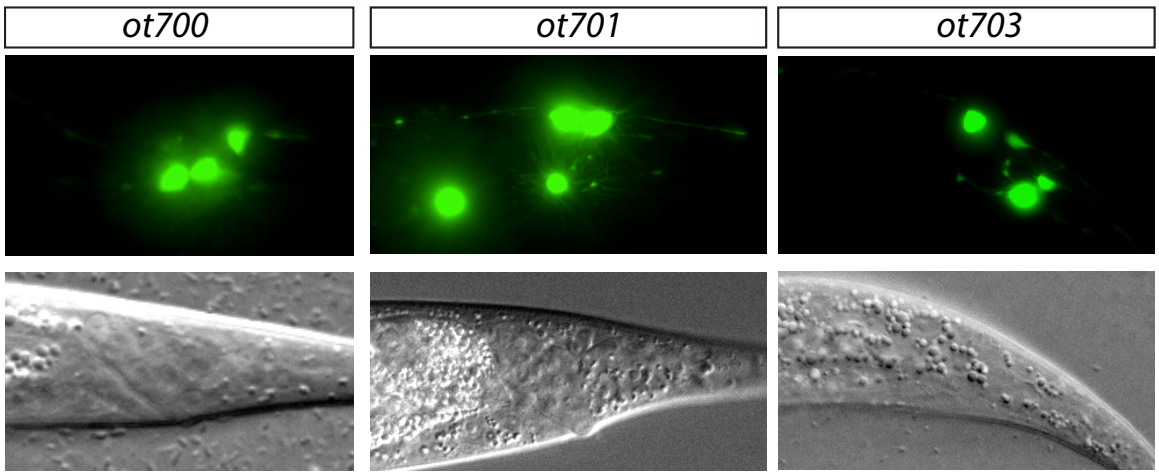


(A) Both *ot699* and *ot702* show extra cells in the anterior half of the animal

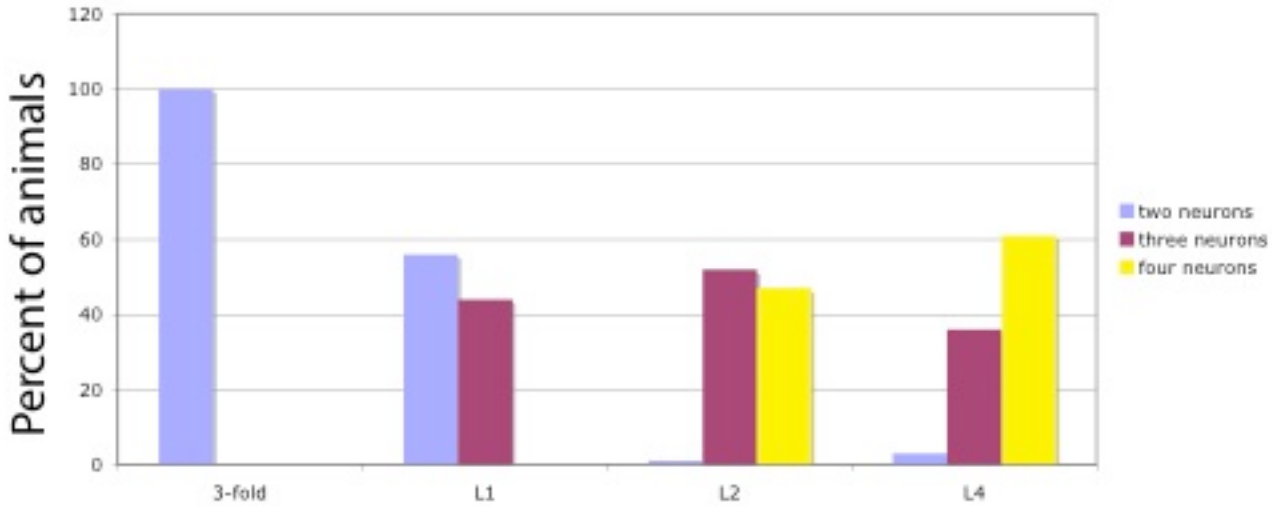
(B) Overlay of *uIs22* (green) and *unc-86::rfp* (red) in *ot702*. The arrow indicates the ectopic cell found in *ot702*.

Figure 3: Ectopic tail expression in Class C mutants

(A)



(B)



(A) Expression of *uIs22* in each of the three Class C mutants

(B) Onset of expression of ectopic cells in *ot701*

Class B mutants had ectopic expression of *mec-3::gfp* in one anterior neuron (Figure 2A). To identify the extra cell, I crossed in *unc-86::rfp*. The extra cell showed overlap with *unc-86* in embryonic and L1 stage, but not by L4 stage (Figure 2B). This change suggests that the extra cell is SDQR, which has *unc-86* expression only in embryonic and early L1 stages. SDQR is the sister cell of AVM, a *mec-3*-expressing mechanosensory neuron.

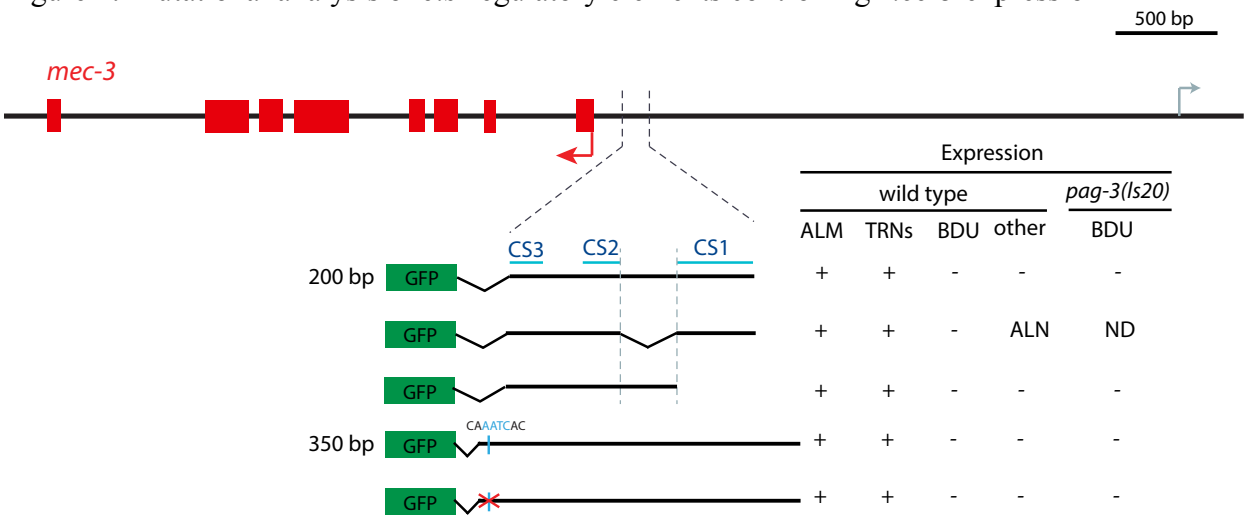
The final class of mutants are those with extra cells in the tail. *mec-3* is normally found in two tail cells, the left and right PLM neurons. In *ot700*, there is expression in one extra cell, while in *ot701* and *ot703* two extra cells show expression (Figure 3A). Identification of the extra cells in *ot701* was enabled in part by looking at onset of expression. In embryonic and L1 stages, there is only expression in two cells in the tail. Beginning in late L1 or L2 stage, extra cells are seen. This late onset suggests the ectopic cells may be postembryonically generated. From the morphology of the axons, I hypothesize they may be PHC neurons.

Cis-*regulatory analysis of the* mec-3 *promoter*

To better understand the mechanism by which PAG-3 represses expression of *mec-3* in BDU neurons, I dissected the *mec-3* promoter to search for motifs necessary for repression. The proximal promoter contains several highly conserved sequences (CS1-CS3), and previous work had shown that mutations in these regions cause ectopic expression in sister cells of the mechanosensory neurons, particularly the PLM sister, ALN (Way et al., 1991). A minimal promoter containing the region from CS1 to CS3 showed expression in ALM as well as the other touch receptor neurons (TRNs) (Figure

4). Deleting the sequence between CS1 and CS2 caused ectopic expression in one pair of cells in the tail, putatively identified by location as the ALN neurons. Neither construct, when crossed into *pag-3*(*ls20*), had ectopic expression in BDU. A conserved putative PAG-3/Senseless/Gfi-1 binding site (Zweidler-Mckay et al., 1996) was identified downstream of CS3. A larger promoter fragment that contained the binding site failed to show expression in BDU either in a wild type background or in *pag-3*(*ls20*), as did a mutation of that binding site.

Figure 4: Mutational analysis of *cis*-regulatory elements controlling *mec-3* expression



Vertical blue bar: predicted PAG-3/Senseless/Gfi binding motif (Zweidler-Mckay et al., 1996). Mutation of the central AATC is indicated by the red cross over the bar. The central AATC was mutated to TTTC. Horizontal blue bars: *mec-3* conserved sequences (Way et al., 1991). TRNs: touch receptor neurons. (+): at least 80% of animals showed bright expression. (-): less than 10% of animals showed expression. $n \geq 50$ for each line.

While a forward genetic screen did not identify any new repressors of *mec-3* in BDU, it is notable that none of the mutants showed broad de-repression. Rather, each mutant had ectopic expression in only one class of neuron. This result mirrors the expression patterns seen when dissecting the *mec-3* promoter ((Way et al., 1991), this work), in which several mutations caused ectopic expression in the sister cells of PLM, but none caused wide-spread sister expression. These results support a piecemeal model, in which *mec-3* is regulated by a variety of cell-specific factors rather than one master activator or repressor. It is unclear whether each of those neuron-specific repressive factors are interacting with Wnt signaling. Genetic screens are unlikely to find mutants in the Wnt signaling pathway, as that pathway affects large portions of the embryo and mutants are likely to cause lethality.

Overexpression of unc-86

In Chapter 2, I presented an experiment in which *unc-86* was driven by a heat shock promoter to create overexpression in ALM. After heat shock at embryonic stages and then 24 hours overnight at 25°C, animals showed ectopic expression of BDU genes with no change in expression of ALM genes. Heat shock creates high levels of *unc-86*, and intriguingly, other attempts at *unc-86* overexpression failed to cause any change in BDU reporters (Table 2). One possibility for this difference in phenotype is that injections of either the *unc-86* genomic locus or the *unc-86* fosmid failed to induce high enough expression levels to allow a dual fate.

Table 2: Overexpression of *unc-86*

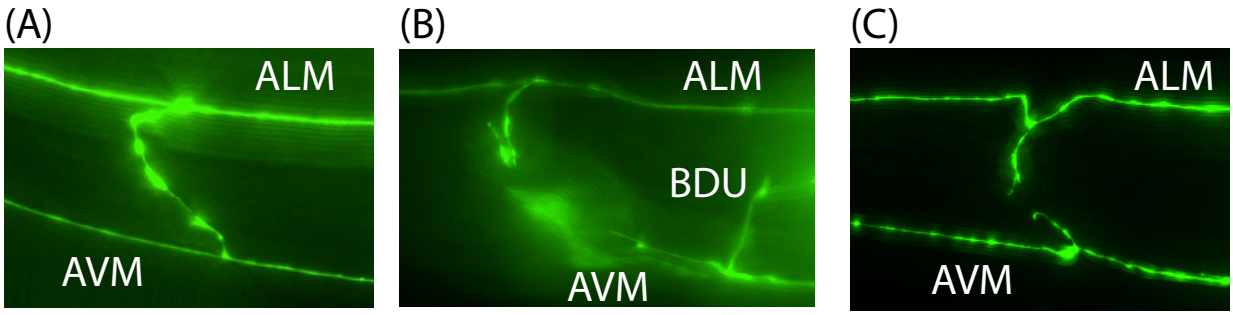
Overexpression construct	Injection concentration	Phenotype	Lines
unc-86 genomic locus	5 ng/ul, complex injection	No change	3/3
	10 ng/ul, complex injection	No change	3/3
	50 ng/ul, simple injection	No change	3/3
unc-86 fosmid	10 ng/ul, complex injection	No change	3/3
	15 ng/ul, complex injection	No change	3/3
	20 ng/ul, complex injection	No change	3/3
hsp::unc-86	3 ng/ul, complex injection	Ectopic expression	2/2

Each construct was injected into an integrated BDU reporter construct. The *unc-86* genomic locus is 9.1kb, including the entire 5.2kb upstream intergenic region and 560 bp downstream of the stop codon. The *unc-86* fosmid is WRM0612cF07. Phenotype was measured by examining expression of the BDU reporter. "No change" indicates no ectopic expression in ALM, while "ectopic expression" indicates expression in ALM.

Loss of BDU-regulating genes causes AVM branching errors

Besides its involvement in the harsh touch circuitry, little is known about the function of BDU. Some evidence, however, has pointed to a role for BDU in physically guiding the branching of AVM neurons into the nerve ring (Walthall and Chalfie, 1988). AVM guidance is importance for its mechanosensory role, as shown by laser ablation studies: in the absence of ALM neurons, AVM neurons can provide partial anterior touch sensitivity. Ablation of both ALM and BDU, however, totally eliminates anterior gentle touch response (Walthall and Chalfie, 1988). Those observations were made before the development of GFP markers allowed easy tracing of neurons in the nerve cord. To further understand the role of BDU in AVM guidance, I used a *mec-4::gfp* reporters to examine AVM and ALM. Under wild-type conditions, the AVM axon makes a dorsal turn into the nerve ring, where it encounters the ALM axon (Figure 5A). I quantified AVM guidance defects by counting the number of branches that failed to meet ALM (Table 3). In animals with the *pag-3(ls20)* mutation, one of the two AVM branches failed to intersect with ALM in 28% of animals (Figure 5B). However, all animals had at least one AVM branch successfully connect with ALM. In *ceh-14(ch3)* mutants, 45% of animals showed defects in one branch, while 7% showed defects in both branches (Figure 5C).

Figure 5: BDU regulatory factors affect AVM guidance



AVM and ALM axon extensions are shown in the nerve ring with *mec-4::gfp* in wild-type (A), *pag-3(ls20)* (B), and *ceh-14(ch3)* (C). In (B), *mec-4::gfp* is also expressed in BDU.

Table 3: AVM guidance defects in *pag-3* and *ceh-14* mutants

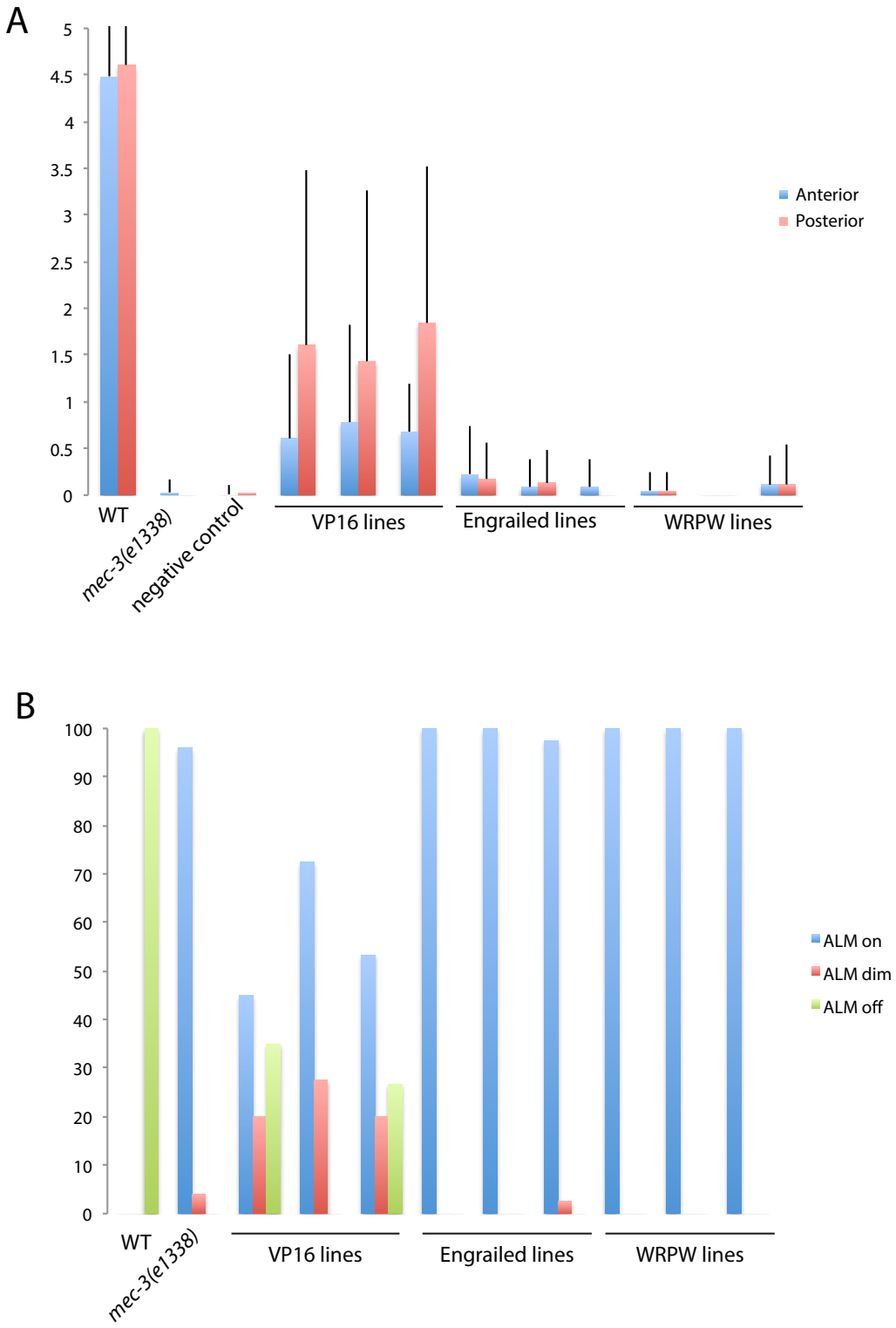
Genotype	0 connections	1 connection	2 connections	n
Wild-type	0%	0%	100%	40
pag-3(ls20)	0%	28%	72%	32
ceh-14(ch3)	7%	45%	48%	29

Guidance defects in AVM are measured by percent of animals in which AVM successfully meets up with ALM in the nerve ring. Number of connections is given out of two, as the AVM axon divides into two branches that each interact with one ALM neuron.

Examination of activation or repression activities of MEC-3

Upon discovering that BDU reporter genes were ectopically expressed in ALM upon loss of *mec-3*, it was not immediately clear how repression of that gene battery occurred. One possibility was that *mec-3* was acting as a transcriptional repressor as well as an activator of the ALM-specific genes; alternatively, it could regulate some other gene that in turn would act as a repressor. To investigate those possibilities, I created several constructs in which the *C-terminal activation domain of* MEC-3 was replaced with either the strong activator VP16, the Engrailed repressor domain, or the repressive WRPW motif. These constructs were all driven by a minimal *mec-3* promoter and used to rescue the mechanosensory defects (Figure 6A) and ectopic BDU marker expression (Figure 6B) of *mec-3*(*e1338*). The VP16 construct was able to partially rescue both aspects of the *mec-3* phenotype, while the Engrailed and WRPW constructs were not able to rescue either. These data support the idea that MEC-3 acts as a transcriptional activator, and it does not directly repress the BDU genes. However, these results are inconsistent with the model I have presented in this thesis, in which MEC-3 sequesters UNC-86 away from the promoters of BDU genes. By that model, all three constructs should have been equally able to rescue the *mec-3* mutant. It is possible that the Engrailed domain or the WRPW motif interfered with MEC-3 interactions with UNC-86. As the results were negative, however, it is also possible that there were problems with the constructs, and *mec-3* was not properly expressed in ALM. Ultimately, I cannot draw any conclusions from these data.

Figure 6: MEC-3 acts as a transcriptional activator



mec-3(e1338) was rescued by a construct consisting of amino acids 1-299 of *mec-3* fused to the strong activator VP16, the strong repressor Engrailed, or nothing (negative control), driven by 350 bp of the *mec-3* promoter.

(A) Constructs were tested for rescue of the *mec-3* mechanosensory defect. Results show average number of responses to light touch on the anterior or posterior of the animal, out of five total touches.

(B) Constructs were tested for rescue of ectopic *flp-10::gfp* expression in ALM.

Materials and Methods

1. Mutant alleles

unc-86(n846), mec-3(e1338), pag-3(ls20), ceh-14(ch3).

2. DNA constructs

All *gfp* constructs for the *cis*-regulatory analysis were generated by subcloning promoter fragments into the pPD95.75 multiple cloning site (Addgene). Linearized plasmids were injected as complex arrays using 5 ng/ul of plasmid, 100 ng/ul of PvuII-digested bacterial genomic DNA, and 5 ng/ul of injection marker. *unc-86* overexpression constructs were injected in a similar manner, using the concentrations indicated in Table 2. Constructs for VP16 and Engrailed experiments were injected at 10 ng/ul of plasmid, 100 ng/ul of PvuII-digested bacterial genomic DNA, and 5 ng/ul of injection marker.

3. Mutagenesis

Animals were mutagenized with EMS according to standard protocols (Brenner, 1974). Three F1 progeny of mutated P0 animals were singled onto individual plates, and their progeny were screened under a dissecting scope equipped with a fluorescent light source.

References

Baumeister, R., Liu, Y., and Ruvkun, G. (1996). Lineage-specific regulators couple cell lineage asymmetry to the transcription of the *Caenorhabditis elegans* POU gene unc-86 during neurogenesis. Genes & Development *10*, 1395–1410.

Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. Genetics *77*, 729-742.

Mizumoto, K., and Sawa, H. (2007). Two βs or not two βs: regulation of asymmetric division by β-catenin. Trends in Cell Biology *17*, 465–473.

Shaham, S., and Bargmann, C.I. (2002). Control of neuronal subtype identity by the *C. elegans* ARID protein CFI-1. Genes & Development *16*, 972–983.

Walthall, W.W., and Chalfie, M. (1988). Cell-cell interactionsi n the guidance of late-developing neurons in *Caenorhabditis elegans*. Science *239*, 643-645.

Way, J.C., Wang, L., Run, J.Q., and Wang, A. (1991). The mec-3 gene contains cis-acting elements mediating positive and negative regulation in cells produced by asymmetric cell division in Caenorhabditis elegans. Genes & Development *5*, 2199–2211.

White, J.G., Southgate, E., Thomson, J.N., and Brenner, S. (1986). The Structure of the Nervous System of the Nematode *Caenorhabditis elegans*. Philosophical Transactions of the Royal Society of London B. Biological Sciences *314*, 1–340.

Zhang, F., Bhattacharya, A., Nelson, J.C., Abe, N., Gordon, P., Lloret-Fernandez, C., Maicas, M., Flames, N., Mann, R.S., Colón-Ramos, D.A., et al. (2014). The LIM and POU homeobox genes *ttx-3* and *unc-86* act as terminal selectors in distinct cholinergic and serotonergic neuron types. Development *141*, 422–435.

Zweidler-Mckay, P.A., Grimes, H.L., Flubacher, M.M., and Tsichlis, P.N. (1996). Gfi-1 encodes a nuclear zinc finger protein that binds DNA and functions as a transcriptional repressor. Molecular and Cellular Biology *16*, 4024–4034.

APPENDIX C: The LIM and POU homeobox genes *ttx-3* and *unc-86* act as terminal selectors in distinct cholinergic and serotonergic neuron types

Feifan Zhang, Abhishek Bhattacharya, Jessica C. Nelson, Namiko Abe, Patricia Gordon, Carla Lloret-Fernandez, Miren Maicas, Nuria Flames, Richard S. Mann, Daniel A. Colón-Ramos and Oliver Hobert. Development 141: 422-435, January 1 2014.

While the terminal selector theory describes how a single transcription factor can direct cell fate by binding to a common conserved DNA binding motif, it does not address the question of how one factor can differentially function in different cell types. This work examines this question as it pertains to two transcription factors: the LIM homeodomain transcription factor TTX-3 and the POU homeodomain transcription factor UNC-86. Here, we find that cell fate is determined in a combinatorial manner, with each transcription factor interacting with distinct cofactors in different cell types.

Much of my work that is presented here was done during the beginning stages of my project, when I was first examining the means by which *unc-86* regulates distinct cell types. I performed genetic tests on markers of IL2, as seen in Figure 8.

© 2014. Published by The Company of Biologists Ltd | Development (2014) 141, 422-435 doi:10.1242/dev.099721

RESEARCH ARTICLE

The LIM and POU homeobox genes *ttx-3* and *unc-86* act as terminal selectors in distinct cholinergic and serotonergic neuron types

Feifan Zhang[1], Abhishek Bhattacharya[1], Jessica C. Nelson[2], Namiko Abe[1], Patricia Gordon[1], Carla Lloret-Fernandez[3], Miren Maicas[3], Nuria Flames[1,3], Richard S. Mann[1], Daniel A. Colón-Ramos[2] and Oliver Hobert[1,4,*]

ABSTRACT

Transcription factors that drive neuron type-specific terminal differentiation programs in the developing nervous system are often expressed in several distinct neuronal cell types, but to what extent they have similar or distinct activities in individual neuronal cell types is generally not well explored. We investigate this problem using, as a starting point, the *C. elegans* LIM homeodomain transcription factor *ttx-3*, which acts as a terminal selector to drive the terminal differentiation program of the cholinergic AIY interneuron class. Using a panel of different terminal differentiation markers, including neurotransmitter synthesizing enzymes, neurotransmitter receptors and neuropeptides, we show that *ttx-3* also controls the terminal differentiation program of two additional, distinct neuron types, namely the cholinergic AIA interneurons and the serotonergic NSM neurons. We show that the type of differentiation program that is controlled by *ttx-3* in different neuron types is specified by a distinct set of collaborating transcription factors. One of the collaborating transcription factors is the POU homeobox gene *unc-86*, which collaborates with *ttx-3* to determine the identity of the serotonergic NSM neurons. *unc-86* in turn operates independently of *ttx-3* in the anterior ganglion where it collaborates with the ARID-type transcription factor *cfi-1* to determine the cholinergic identity of the IL2 sensory and URA motor neurons. In conclusion, transcription factors operate as terminal selectors in distinct combinations in different neuron types, defining neuron type-specific identity features.

KEY WORDS: *Caenorhabditis elegans*, Homeobox, Neuron differentiation

INTRODUCTION

The development of the nervous system is a multistep process that employs a series of sequentially acting regulatory factors that successively restrict and determine cellular fates. During the process of terminal differentiation, individual neuron types acquire specific, hard-wired features that are maintained by the neuron type throughout the life of the animal. A number of transcription factors have been identified that initiate and maintain specific terminal differentiation programs in the developing nervous system (Hobert, 2011). For example, in mouse, the Nurr1 (Nr4a2) transcription factor initiates and maintains the terminal differentiation program of dopaminergic neurons in the midbrain (Smidt and Burbach, 2009), whereas the Pet1 transcription factor initiates and maintains the terminal differentiation program of serotonergic neurons (Liu et al., 2010). However, few neuronal transcription factors are expressed exclusively in only one specific neuronal cell type (Gray et al., 2004; Lein et al., 2007). For example, in addition to being expressed in midbrain dopaminergic neurons, Nurr1 is expressed in other non-dopaminergic neuronal cell types in which its function is not well understood, such as the adult olfactory bulb, specific cortical areas and the hippocampus (Zetterström et al., 1996). The expression of a given transcription factor in distinct neuronal populations poses the fundamental question of whether there are underlying common themes in the activity of the transcription factor in distinct neuronal cell types.

We have undertaken a systematic, in-depth comparison of the activity of two transcription factors in the development of several distinct neuronal cell types in the nematode *C. elegans*, examining whether there are indeed conceptual similarities in the activities of a given transcription factor in distinct neuron types. We used, as a starting point, a member of the LIM homeobox gene family, an ancient family of neuronal patterning genes that display complex expression patterns in the nervous system of many different species, from invertebrates to vertebrates (Hobert and Westphal, 2000; Simmons et al., 2012; Srivastava et al., 2010). One unifying theme is their expression in terminally differentiating neurons (Hobert and Ruvkun, 1998; Moreno et al., 2005). We focus here on the *ttx-3* LIM homeobox gene, which is the sole *C. elegans* member of the Lhx2/9 subclass of LIM homeobox genes. In vertebrates, Lhx2 is expressed in multiple neuronal cell types and is required for the differentiation of olfactory sensory neurons (Hirota and Mombaerts, 2004; Kolterud et al., 2004), the specification of cortical neuron fate (Mangale et al., 2008) and the differentiation of thalamic neurons (Peukert et al., 2011). Whether there is a common theme in the function of Lhx2 in these distinct neuronal cell types is not known.

The *C. elegans* Lhx2/9 ortholog *ttx-3* is exclusively expressed in a small number of neurons in distinct head ganglia (Altun-Gultekin et al., 2001). *ttx-3* null animals display broad differentiation defects in the cholinergic AIY interneuron class. AIY interneurons of *ttx-3* null mutants are generated and still express pan-neuronal features, but fail to express scores of terminal identity markers that define the functional properties of AIY, including genes required to synthesize and package acetylcholine, genes encoding neuropeptide receptors, various types of ion channels and many others (Altun-Gultekin et al., 2001; Hobert et al., 1997; Wenick and Hobert, 2004). TTX-3 exerts this control through direct binding to a *cis*-regulatory motif shared by all of its target genes. *ttx-3* expression is turned on in the

[1]Department of Biochemistry and Molecular Biophysics, Columbia University Medical Center, New York, NY 10032, USA. [2]Department of Cell Biology, Yale University School of Medicine, New Haven, CT 06520, USA. [3]Instituto de Biomedicina de Valencia, IBV-CSIC, 46010 Valencia, Spain. [4]Howard Hughes Medical Institute, Columbia University Medical Center, New York, NY 10032, USA.

*Author for correspondence (or38@columbia.edu)

Received 31 May 2013; Accepted 11 October 2013

Development (2014) doi:10.1242/dev.099721

neuroblast that generates AIY and its expression is maintained throughout the life of the neuron through an autoregulatory feedback loop (Bertrand and Hobert, 2009) to ensure persistent expression of its target genes. A number of transcription factors have been described in the *C. elegans* nervous system that display similar broad-ranging effects on the terminal differentiation programs executed by the neurons in which they are expressed. These transcription factors have been called 'terminal selectors' (Hobert, 2008; Hobert, 2011). It is still an open question how broadly the terminal selector concept applies throughout the nervous system; that is, how common it is that many distinct and functionally unrelated identity features of a specific neuron type are directly co-regulated by a transcription factor or a combination of transcription factors.

Here, we investigate the role of *ttx-3* in two additional neuron classes in which it is normally expressed, namely the cholinergic AIA interneuron class and the serotonergic NSM neuron class. We find in all three neuron classes that there is a common theme of *ttx-3* function in that it is broadly required to induce many distinct and functionally unrelated terminal identity features of the respective neuron class. Yet the downstream targets of *ttx-3* in these neuron classes are distinct and are determined by the cooperation of *ttx-3* with a distinct set of transcription factors in different neuron classes. One of these factors is the POU homeobox gene *unc-86*, which is required together with *ttx-3* to control the identity of the serotonergic NSM neurons. *unc-86* in turn cooperates with the ARID-type transcription factor *cfi-1* to control many terminal identity features of the cholinergic IL2 sensory and URA motor neurons. Our studies therefore provide further support for the terminal selector concept and show that, in combination with other regulatory factors, one factor can serve as terminal selector in distinct neuronal cell types regulating distinct neuronal differentiation programs.

RESULTS

Expression pattern of *ttx-3* in the *C. elegans* larval and adult nervous system

A *ttx-3* reporter gene that contains the *ttx-3* locus together with a few kilobases upstream but no downstream sequences (*ttx-3*^{promA}*::gfp*; Fig. 1) was previously shown to be continuously expressed in five distinct neuronal cell types: the cholinergic AIY and AIA interneuron classes, the ASI and ADL chemosensory neuron classes and a previously uncharacterized neuronal pair in the pharyngeal nervous system (Altun-Gultekin et al., 2001). Transient expression was observed in the AIN and SMDD neurons at embryonic stages (Bertrand and Hobert, 2009). A fosmid reporter construct, which contains more than 30 kb surrounding the *ttx-3* locus and which

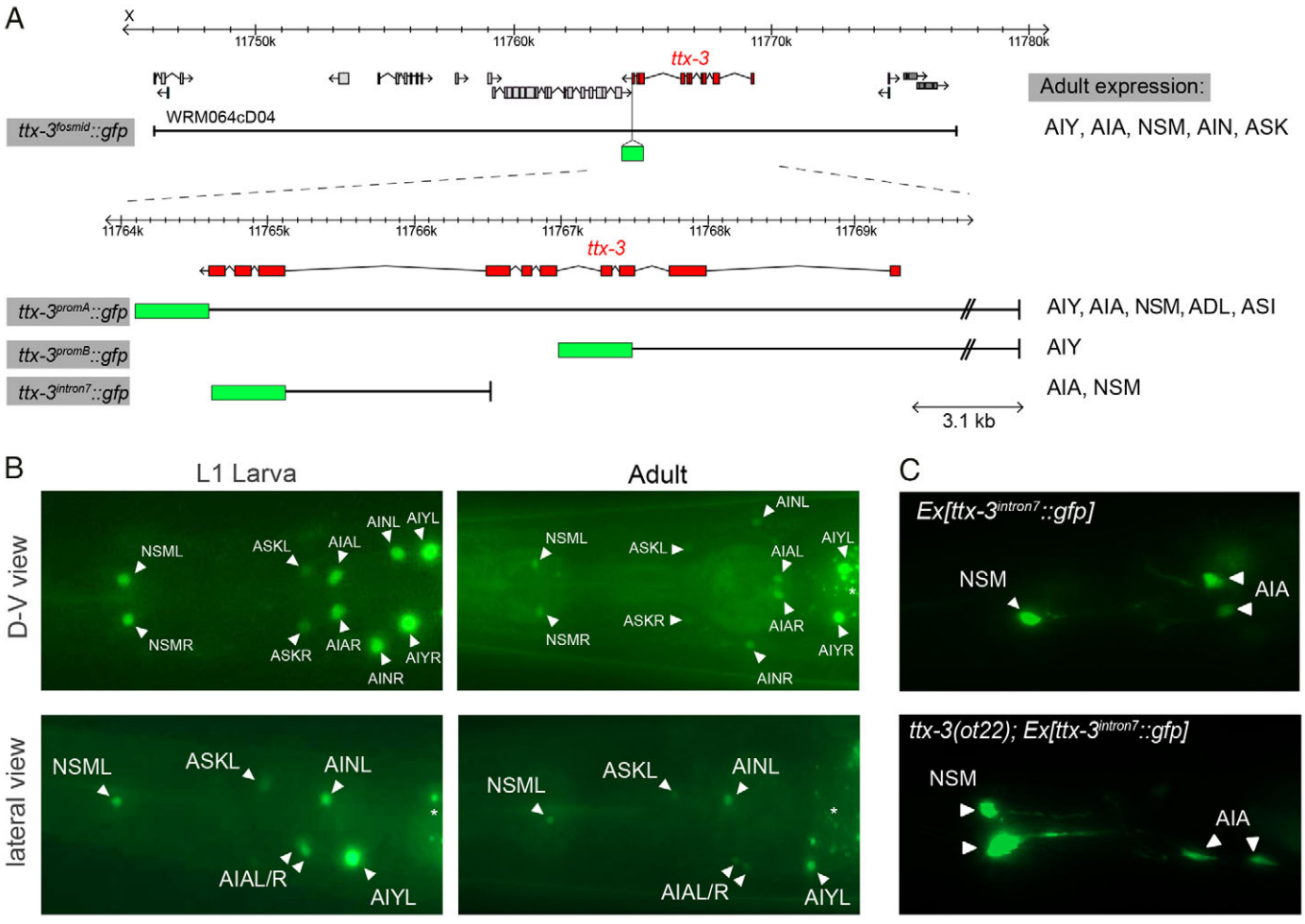


Fig. 1. Expression pattern of the *C. elegans ttx-3* LIM homeobox gene. (A) *ttx-3* expression constructs and summary of neuronal expression pattern. The *promA::gfp* and *promB::gfp* constructs were described previously (Altun-Gultekin et al., 2001; Wenick and Hobert, 2004) and are shown here for comparison only. (B) *ttx-3* fosmid expression (*wgIs68*) in first larval stage animals and in adult animals. D-V, dorsoventral. White asterisks indicate gut autofluorescence. (C) The seventh intron of the *ttx-3* locus contains *cis*-regulatory elements driving reporter gene expression in AIA and NSM neurons. These regulatory elements do not depend on *ttx-3*. Expression is shown in adult animals.

Development (2014) doi:10.1242/dev.099721

rescues the AIY differentiation defect of *ttx-3* mutant animals, mirrors the expression of the smaller, locus-restricted reporter construct in the AIY, the AIA, the AIN and the pharyngeal neuron class (Fig. 1). Based on position, morphology and colabeling with the NSM marker *mgl-1::mCherry*, we identified the pharyngeal neurons that express *ttx-3* as the NSM neuron pair. The NSM neurons are serotonergic, neurosecretory cells that are thought to be involved in sensing food (Albertson and Thomson, 1976; Harris et al., 2011; Horvitz et al., 1982).

There are also notable differences in the expression pattern of the fosmid reporter and the smaller reporters. First, expression in the AIN neurons is maintained throughout development with the fosmid reporter, whereas it is restricted to embryos with smaller reporters (Bertrand and Hobert, 2009). Second, the expression in amphid sensory neurons is markedly different. In larval and adult animals, the fosmid reporter is expressed in the ASK neuron class, whereas the smaller reporters are expressed in the ADL and ASI sensory neurons (Fig. 1).

Previous studies have shown that *ttx-3* expression in the AIY interneuron pair is controlled by a distal initiator element ~1 kb upstream of the *ttx-3* locus and a maintenance element in the second intron of the *ttx-3* locus (Bertrand and Hobert, 2009; Wenick and Hobert, 2004). We find that the expression of *ttx-3* in the NSM and AIA is controlled via regulatory elements present in the seventh intron of the *ttx-3* locus (Fig. 1). As mentioned above, *ttx-3* expression is maintained throughout the life of the AIA and NSM neurons, but maintained expression of a *ttx-3* reporter gene construct (*ttx-3*^{intron7}*::gfp*; Fig. 1A) in the AIA and NSM neuron types does not require *ttx-3* gene activity (Fig. 1C).

ttx-3 controls the differentiation program of AIA interneurons

We focused our analysis of *ttx-3* mutants on the cholinergic AIA interneurons and the serotonergic NSM neurons, which both continuously express *ttx-3* throughout their lifetime. We have previously reported that expression of the marker of cholinergic identity, *unc-17* (vesicular ACh transporter), as well as the expression of an orphan G protein-coupled receptor (GPCR), *sra-11*, is reduced in the AIA neurons of *ttx-3* mutants (Altun-Gultekin et al., 2001). We extended this analysis by examining the expression of seven additional markers of terminal AIA fate: the choline reuptake transporter encoded by *cho-1*; the metabotropic glutamate receptor *mgl-1*; the ionotropic glutamate receptor *glr-2*; the neuropeptides *flp-2* and *ins-1*; the receptor tyrosine kinase *scd-2*; and the receptor guanylyl cyclase *gcy-28d.* Each of these markers is expressed in terminally differentiated AIA interneurons and several of them have previously been implicated in AIA interneuron function (Shinkai et al., 2011; Tomioka et al., 2006). The expression of each of these seven markers is affected in the AIA neurons of *ttx-3* mutants (Fig. 2). Their expression in other neuron types is unaffected in *ttx-3* mutants, with the exception of two markers that are also downregulated in NSM neurons (*mgl-1*, *scd-2*, as described below). *ttx-3* is likely to act cell-autonomously since the AIA differentiation defects are rescued in transgenic *ttx-3* mutant animals that express *ttx-3* cDNA under control of the *ins-1* promoter (supplementary material Table S1).

AIA neurons remain present in the *ttx-3* null mutant, as assessed by the weak but recognizable expression of some terminal differentiation genes (Fig. 2). However, their normally unipolar neurite morphology appears disrupted; ectopic branches can be observed to emanate from the cell body and the main neurite appears blebbed in *ttx-3* mutants (supplementary material Fig. S2). The AIY interneurons, which have a unipolar axon morphology similar to that of AIA interneurons in wild-type animals, display similar morphological defects in *ttx-3* mutants (Hobert et al., 1997). The expression of terminal identity markers that label several distinct neuron types that are lineally related to AIA is not altered (data not shown) (Altun-Gultekin et al., 2001), suggesting that the AIA neuron pair might remain in an undifferentiated state, rather than switching to an alternate fate. Based on a more extensive cell fate marker analysis, a similar conclusion was previously drawn about the fate of the AIY neuron class in *ttx-3* mutants (Altun-Gultekin et al., 2001). Taken together, our fate marker and morphological analyses indicate that *ttx-3* broadly affects the AIA terminal differentiation program. These effects are comparable to the previously described broad effects that loss of *ttx-3* has on the terminal differentiation of AIY interneurons.

A shared *cis*-regulatory signature of AIA-expressed terminal identity features

On a mechanistic level, *ttx-3* operates in a distinct manner in the AIA versus AIY neurons since it operates with distinct co-factors and through distinct *cis*-regulatory elements. The co-factor of *ttx-3* in AIY, the *ceh-10* homeobox gene (Altun-Gultekin et al., 2001), is not expressed in AIA neurons, and AIA neurons display no differentiation defects in *ceh-10* null mutants (two markers tested). Moreover, the *cis*-regulatory motifs through which *ttx-3* acts to control AIY versus AIA identity are distinct. In the AIY neurons, *ttx-3* acts on its many target genes through a *cis*-regulatory motif, termed the 'AIY motif', that provides a cooperative binding site for a TTX-3–CEH-10 heterodimer (Wenick and Hobert, 2004). Mutation of the AIY motif in a locus that is expressed in AIY and AIA neurons, the cholinergic *cho-1* locus, results in a severe reduction in expression in the AIY interneurons but not in the AIA interneurons (Fig. 3A).

In the AIA neurons, by contrast, *ttx-3* acts through a distinct *cis*-regulatory signature, which we deciphered through a mutational analysis of the *cis*-regulatory control regions of three AIA-expressed, *ttx-3*-dependent terminal differentiation genes: *mgl-1*, *ins-1* and *cho-1*. We generated transgenic animals that express nested, shorter versions of these three reporters and identified a 259 bp element in the *cho-1* promoter, a 74 bp element in the *mgl-1* promoter and a 68 bp element in the *ins-1* promoter that are sufficient to direct *gfp* expression to AIA neurons (Fig. 3A-C). Examining these elements for common patterns, we noted that all these elements contain a shared and phylogenetically conserved G(A/G)ATC motif (Fig. 3D). Mutating this motif in the context of any of the three promoters resulted in a reduction of AIA expression of the respective reporter (Fig. 3A-C). In the case of *mgl-1*, two G(A/G)ATC motifs are present in the minimal promoter; mutation of either causes an intermediate reduction in reporter gene expression, and mutation of both motifs results in complete loss of expression (Fig. 3A-C).

Since G(A/G)ATC does not match the consensus binding site for a LIM homeodomain transcription factor such as TTX-3, we also examined the minimal reporters for the presence of conserved TAAT motifs, which comprises the core consensus site for LIM homeodomain transcription factors (Berger et al., 2008). We indeed found several TAAT motifs in the three *cis*-regulatory modules and for each of them we identified a TAAT motif that, when mutated, affected reporter gene expression *in vivo* (Fig. 3A-C). These TAAT motifs can be assembled into a larger sequence matrix, TAATTNGA (Fig. 3D). In two cases, mutation of the TAATTNGA alone affected reporter gene expression, whereas in the third case (*cho-1*) a

Development (2014) doi:10.1242/dev.099721

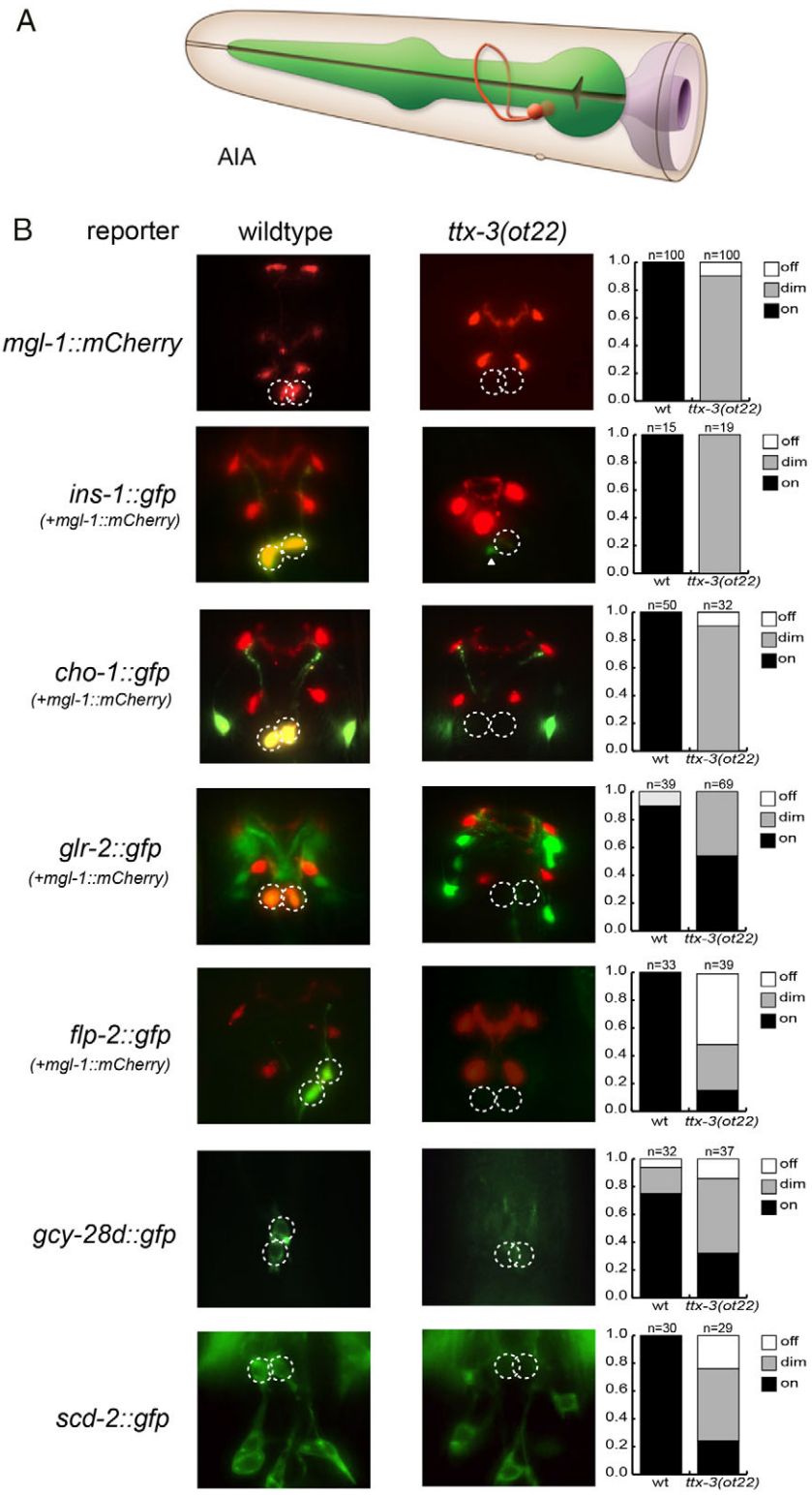


Fig. 2. *ttx-3* affects the terminal differentiation of AIA neurons. (A) Schematic representation of the AIA interneuron pair [reproduced with permission (Altun et al., 2002-2013)]. (B) The expression of terminal differentiation markers of AIA identity is affected in *ttx-3* mutants. Reporter gene arrays were crossed into *ttx-3(ot22)* null mutants. Positions of AIA neurons are outlined (dashed circles). The fraction of animals that show the indicated phenotype is presented in the bar charts. Transgenic arrays are: *otIs317* for *mgl-1*, *otIs326* for *ins-1*, *otIs379* for *cho-1*, *otEx4687* for *glr-2* and *otEx5056* for *flp-2* (see Materials and methods for more detail on the arrays; the *Ex[gcy-28d::gfp]* and *Ex[scd-2::gfp]* arrays were kindly provided by Takeshi Ishihara). Anterior is up in all panels.

complete loss of expression can only be observed upon simultaneous mutation of both the GAATC motif and the TAAT-containing motif (Fig. 3A). The residual AIA expression of a *cho-1* reporter construct in which the GAATC motif is mutated, but the TAAT motif is left intact, is abolished in *ttx-3* mutants (data not shown), consistent with *ttx-3* operating through the TAAT motif.

We examined whether the TAATTNGA motif is indeed a TTX-3 binding site using gel shift assays with bacterially produced TTX-3 protein and probes derived from the *mgl-1* and *cho-1* locus. We found that TTX-3 is able to bind these sites *in vitro* (Fig. 3E). Deletion of the TAAT site that is required for reporter gene expression *in vivo* resulted in the loss of TTX-3 binding *in vitro* (Fig. 3E).

The combination of G(A/G)ATC and TAAT motifs might define a *cis*-regulatory signature that is generally required for gene expression in AIA neurons, since we found a combination of these two motifs to be present in the *cis*-regulatory control regions of the

Development (2014) doi:10.1242/dev.099721

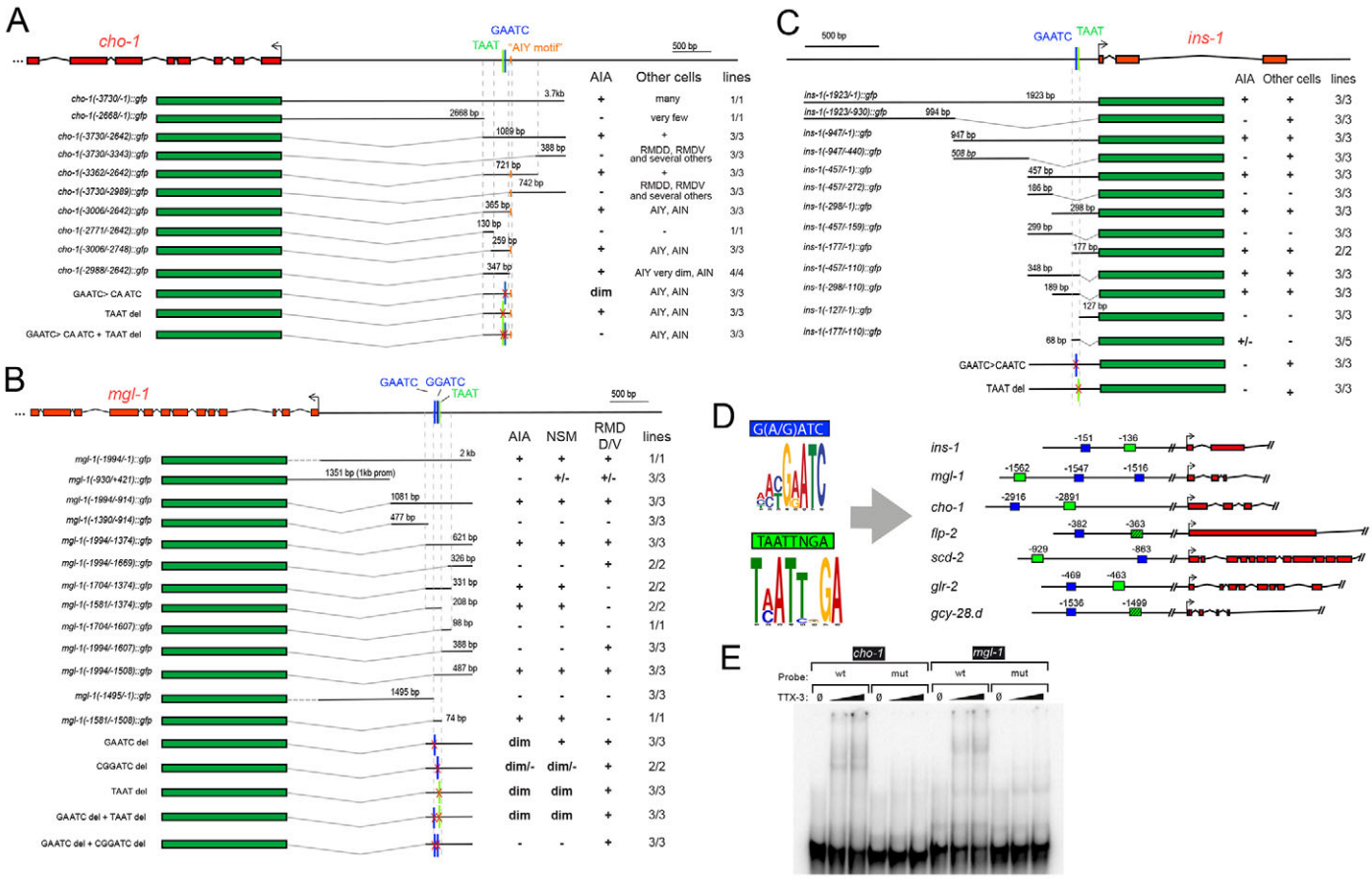


Fig. 3. Co-regulation of AIA-expressed genes by two *cis*-regulatory motifs. (A-C) Mutational dissection of the *cis*-regulatory elements of three AIA-expressed terminal identity markers. (D) Position weight matrix of the two motifs required for AIA expression, based on the motifs from *ins-1*, *cho-1* and *mgl-1* and orthologs in other nematode species. Perfect (filled box) and imperfect (stippled box) matches to the two *cis*-regulatory motifs [blue, G(A/G)ATC; green, TAATTNGA] in other AIA terminal identity markers are shown on the right. (E) TTX-3 binds to *cho-1* and *mgl-1* regulatory elements containing the HD (TAAT) motif. Deletion of the HD motif abolishes binding. EMSA was performed with 250 nM and 750 nM TTX-3.

other four *ttx-3*-dependent terminal AIA markers (Fig. 3D). Taken together, these data show that AIA identity features are co-regulated by a shared *cis*-regulatory signature that is controlled by TTX-3 and an as yet unknown co-factor.

ttx-3 controls the terminal differentiation of serotonergic NSM neurons

We next analyzed the effect of loss of *ttx-3* on the terminal differentiation program of the serotonergic NSM neurons, a neuron type that has not previously been examined in *ttx-3* mutants. Many terminal identity markers of NSM have been described, including the battery of genes that are required to synthesize, package and reuptake serotonin: *tph-1/TPH* (tryptophan hydroxylase), *cat-4/GTPCH* (GTP cyclohydrolase), *cat-1/VMAT* (vesicular monoamine transporter), *bas-1/AAAD* (aromatic amino acid decarboxylase) and *mod-5/SERT* (serotonin reuptake transporter) (Fig. 4A) (Jafari et al., 2011; Ranganathan et al., 2001; Sze et al., 2002). Previous expression analysis of a vesicular glutamate transporter, *eat-4*, suggested that NSM might use the neurotransmitter glutamate (Lee et al., 1999). However, a fosmid-based *eat-4* reporter does not show expression in NSM neurons (Serrano-Saiz et al., 2013) (supplementary material Fig. S1A).

To broaden the spectrum of available terminal markers, we analyzed the expression of other *C. elegans* orthologs of enzymes involved in monoaminergic transmitter metabolism (Fig. 4A) and identified another NSM-expressed terminal marker, *ptps-1* (Fig. 4C; supplementary material Fig. S3). In addition to examining these serotonin (5HT)-related markers, we also examined the expression of three metabotropic neurotransmitter receptors (*mgl-1*, *mgl-3*, *dop-3*), three neuropeptides (*nlp-13*, *flp-4*, *nlp-3*), a glycoprotein hormone alpha subunit (*flr-2*) and a receptor tyrosine kinase (*scd-2*). All of these genes are expressed throughout the life of the NSM neurons. As mentioned above, *scd-2* and *mgl-1* are also expressed in AIA neurons, where their expression is affected by *ttx-3*. We find that the expression of five of these 14 NSM terminal identity markers is either partially or completely eliminated in the NSM neurons of *ttx-3* null mutants (Fig. 4C, Fig. 5, Table 1). *ttx-3* is likely to act cell-autonomously since we can rescue the NSM differentiation defects by driving *ttx-3* cDNA under the control of a *cat-1* promoter fragment, which is expressed in a subset of monoaminergic neurons of *C. elegans* (supplementary material Table S1).

The POU homeobox gene *unc-86* also controls NSM identity

We recently reported that the effects of the loss of a terminal selector type transcription factor in dopaminergic neurons can be partially compensated for by other, co-expressed terminal selectors (Doitsidou et al., 2013). Therefore, we considered the possibility that

Development (2014) doi:10.1242/dev.099721

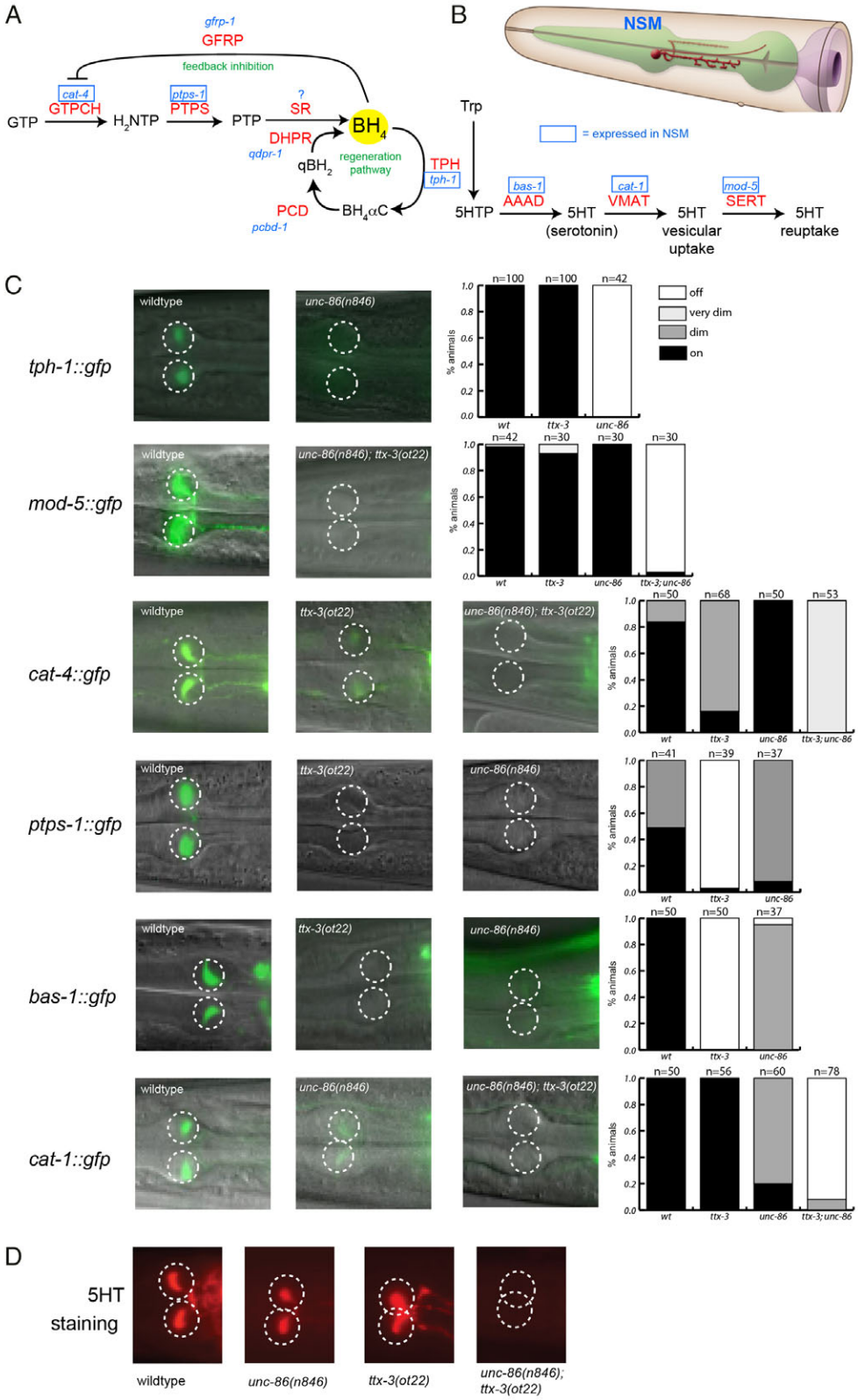


Fig. 4. The effect of *unc-86* and *ttx-3* on the serotonergic identity of NSM neurons. (A) The 5HT pathway including tetrahydrobiopterin biosynthesis genes (Deneris and Wyler, 2012). '?' indicates that a unique homolog of SR could not be identified in the worm genome. (B) Schematic representation of the NSM interneuron pair [reproduced with permission (Altun et al., 2002-2013)]. (C) The expression of serotonergic identity features of NSM (dashed circles) is affected in *unc-86(n846)*, *ttx-3(ot22)* or *unc-86(n846); ttx-3(ot22)* double-null mutants. Reporter gene arrays were crossed into the respective mutant backgrounds. Transgenic arrays are: *zdIs13* for *tph-1*; *otEx4781* for *mod-5*; *otIs225* for *cat-4*; *otEx5280* for *ptps-1*; *otIs226* for *bas-1*; and *otIs224* for *cat-1* (see Materials and methods for more detail on the arrays). Images are only shown for mutant genotypes with effects on reporter expression. (D) Serotonin antibody staining. Thirty animals were scored for each genotype. In the double mutant, no animal showed staining in NSM (circled), whereas in the other genotypes all animals showed staining.

Development (2014) doi:10.1242/dev.099721

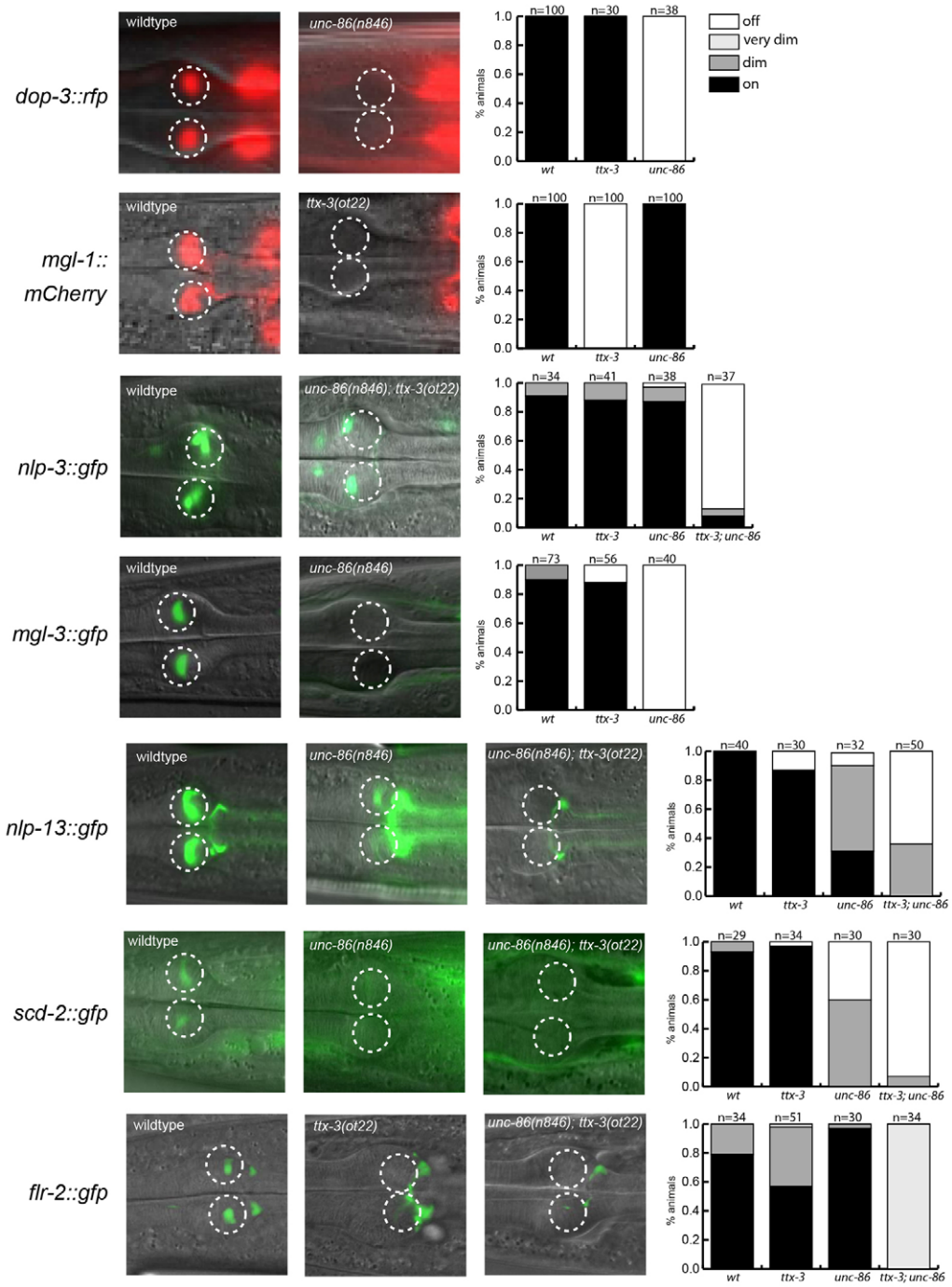


Fig. 5. The effect of *unc-86* and *ttx-3* on other identity features of NSM neurons. The expression of other identity features of NSM is also affected in *unc-86(n846)*, *ttx-3(ot22)* or *unc-86(n846); ttx-3(ot22)* double-null mutants. Reporter gene arrays were crossed into the respective mutant backgrounds. Transgenic arrays are: *vsIs33* for *dop-3*; *otIs317* for *mgl-1*; *otEx5163* for *nlp-3*; *otEx5364* for *mgl-3*; *otEx5163* for *nlp-13*; *otEx5055* for *scd-2*; and *otEx5363* for *flr-2* (see Materials and methods for more detail on the arrays). Micrographs are only shown for mutant genotypes with effects on reporter expression. Dashed circles indicate the position of NSM neurons. See also Table 1.

the lack of an impact of *ttx-3* loss on nine out of 14 NSM markers could be due to the activity of compensatory terminal selector type transcription factors. We sought to identify such a factor, focusing on two homeodomain transcription factors previously shown to be expressed in NSM, namely the empty spiracles homolog *ceh-2* and the POU homeobox gene *unc-86* (Aspöck et al., 2003; Finney and Ruvkun, 1990). We observed no NSM differentiation defects in *ceh-2* null animals (data not shown), but we observed striking NSM differentiation defects in *unc-86* mutants. Loss of *unc-86* was previously shown to affect the expression of *tph-1* and *cat-1* in NSM neurons, but without effect on 5HT antibody staining (Sze et al., 2002). Other differentiation features of NSM neurons had not previously been examined in *unc-86* mutants. Upon examining the expression of all 14 markers of NSM fate in *unc-86* null mutants, we found that the expression of eight is partially or completely eliminated (Figs 4, 5, Table 1).

To examine whether *unc-86* directly affects the expression of these terminal identity features, we analyzed the *cis*-regulatory control regions of four of them: *tph-1*, *bas-1*, *cat-1* and *cat-4*. Through mutational analysis, we defined small (~200 bp) elements that still yielded expression in the NSM neurons (Fig. 6) and, within each of these elements, we identified predicted POU homeodomain binding sites (Rhee et al., 1998). We introduced mutations into these sites in the context of two loci (*tph-1* and *bas-1*) and found that these mutations resulted in a loss of reporter gene expression *in vivo* (Fig. 6A,B). Gel shift analysis further confirmed that these POU homeodomain sites indeed bind bacterially produced UNC-86 protein *in vitro* (Fig. 6E).

Development (2014) doi:10.1242/dev.099721

Table 1. Summary of the effects of *ttx-3* and *unc-86* null mutants on terminal NSM identity markers

Identity feature	Function	*ttx-3(–)*	*unc-86(–)*	*unc-86(–); ttx-3(–)*	Interaction
cat-1	5HT pathway	wt	dim	off	Synergism
cat-4	5HT pathway	dim	wt	very dim	Synergism
mod-5	5HT pathway	wt	wt	off	Synergism
nlp-13	Neuropeptide	wt	dim	off	Synergism
nlp-3	Neuropeptide	wt	wt	off	Synergism
flr-2	Transmembrane	wt	wt	dimmer	Synergism
scd-2	Kinase	wt	dim	off	Synergism
flp-4	Neuropeptide	dim	wt	stronger expression	Synergism
5HT antibody staining	Neurotransmitter	wt	wt	off	Synergism
bas-1	5HT pathway	off	dim	n.d.	?
ptps-1	5HT pathway	off	dim	n.d.	?
tph-1	5HT pathway	wt	off	n.d.	?
mgl-3	GPCR	wt	off	n.d.	?
dop-3	GPCR	wt	off	n.d.	?
mgl-1	GPCR	off	wt	n.d.	?

gfp reporter (or antibody staining): wt, as bright as in wild-type animals; dim, dimmer than in wild type; off, no expression observed.
n.d., not determined because single mutant already shows completely penetrant loss of expression.
Gray shading indicates presence of defect.

unc-86 cooperates with *ttx-3* to control NSM identity

We noted that terminal markers of NSM identity that were severely affected in *unc-86* mutants tended to be those that were weakly or unaffected in *ttx-3* mutants; vice versa, markers unaffected in *ttx-3* mutants tended to be affected in *unc-86* mutants (Table 1). Even though this observation might simply indicate that *unc-86* and *ttx-3* act completely independently of one another, we considered the possibility that *unc-86* and *ttx-3* might collaboratively control NSM identity but that their relative importance may be distinct for different target genes. To investigate this possibility, we examined *unc-86; ttx-3* double-null mutants and found that markers that are either partially or unaffected in *ttx-3* and *unc-86* single mutants are more strongly affected in the double mutant (Figs 4, 5, Table 1). This also holds for 5HT antibody staining, which is not affected in either single mutant but completely abrogated in the *ttx-3*; *unc-86* double mutant (Fig. 4D), probably owing to the combined effect that both genes have on the expression of the 5HT reuptake transporter *mod-5*. As summarized in Table 1, nine of the 15 tested identity features (14 reporter genes and 5HT antibody staining) are affected by both *ttx-3* and *unc-86*, with effects either visible in both single mutants, or as a non-additive, synergistic effect revealed in the double mutant. As described below, there are also synergistic effects of *ttx-3* and *unc-86* on NSM morphology. In six of the 15 cases, either *ttx-3* or *unc-86* already has completely penetrant effects (Table 1). Taken together, these data argue that *unc-86* and *ttx-3* jointly control terminal NSM differentiation. The mechanistic basis of the cooperation is unclear at present because we have so far not been able to identify functional TTX-3 binding sites in terminal NSM identity marker genes.

To further examine potential interactions of *unc-86* and *ttx-3*, we investigated whether they affect each others expression. We find that continuous expression of *unc-86* in NSM neurons depends on *unc-86* itself [autoregulation of *unc-86* was also previously noted (Baumeister et al., 1996)], but not on *ttx-3* (supplementary material Fig. S1B,C). Vice versa, *ttx-3* expression in NSM neurons is not affected in *unc-86* or in *unc-86; ttx-3* mutants (data not shown).

unc-86 and *ttx-3* affect axonal arborization and presynaptic specializations

Apart from affecting the expression of terminal identity markers, loss of *unc-86* and *ttx-3* also results in specific effects on the morphology of the NSM neurons. During embryonic stages, these neurons normally extend a neurite posteriorly toward the nerve ring, which then bifurcates to form a ventral and a dorsal neurite (Axäng et al., 2008) (Fig. 4B). We observe that in *unc-86(n846)* mutants the NSM somas are correctly positioned but there are significant defects in outgrowth of the ventral neurite (61% of animals show outgrowth defects; n=31). By contrast, in *ttx-3(ot22)* mutants, the primary defect observed in ventral neurite outgrowth is the formation of aberrant bifurcations (41% of animals show such defects; n=39). Both *ttx-3* and *unc-86* single mutants also show defects in dorsal axon termination [25% (n=32) of *ttx-3* mutants and 29% (n=34) of *unc-86* mutants].

In early larval stages, ventral NSM neurites begin to extend elaborate arbor structures onto the nerve ring target field (Axäng et al., 2008). These axon arborizations require the NSM-expressed netrin receptor UNC-40/DCC, which is tightly localized to puncta within the main shaft of the NSM neurite and at the tips of axon arbors (Nelson and Colón-Ramos, 2013). These arbor structures persist into the adult stage and contain presynaptic sites, as assessed with a *rab-3* marker (Nelson and Colón-Ramos, 2013). In *ttx-3* mutants, these ultrastructural features are unaffected, but *unc-86* mutants display a highly penetrant defect in axon arborization (Fig. 7A,C). Furthermore, *unc-86* mutants display defects in the dynamic regulation of UNC-40 localization (Fig. 7B,D). In wild-type animals, UNC-40::GFP is diffusely distributed at the L1 stage and becomes localized to bright puncta in the NSM neurite and at the tips of axon arbors as axons are arborizing at the L4 stage. By the adult stage, UNC-40::GFP again becomes diffusely distributed. However, a significant fraction of *unc-86* mutant NSMs retain a juvenile-like pattern of UNC-40 localization during the adult stage, in which UNC-40::GFP remains localized to bright puncta (Fig. 7B).

We observe synergistic morphological defects in *unc-86(n846); ttx-3(ot22)* double mutants. Ventral neurites never reach the middle of the pharyngeal isthmus and are often truncated immediately following the guidance decision to turn posteriorly (100% premature ventral neurite termination, n=18; Fig. 7E). Furthermore, neurites contain large anterior swellings not seen in wild-type animals (33% contain additional anterior swellings, n=18; Fig. 7E). The morphological appearance of NSM neurites in *unc-86; ttx-3* mutants is reminiscent of the normal morphology of M3 neurons (Albertson and Thomson, 1976), which are lineally related to NSM (Sulston et al., 1983). M3 neurons are glutamatergic (Lee et al., 1999) and we indeed find that in *unc-86* mutants the vesicular glutamate

Development (2014) doi:10.1242/dev.099721

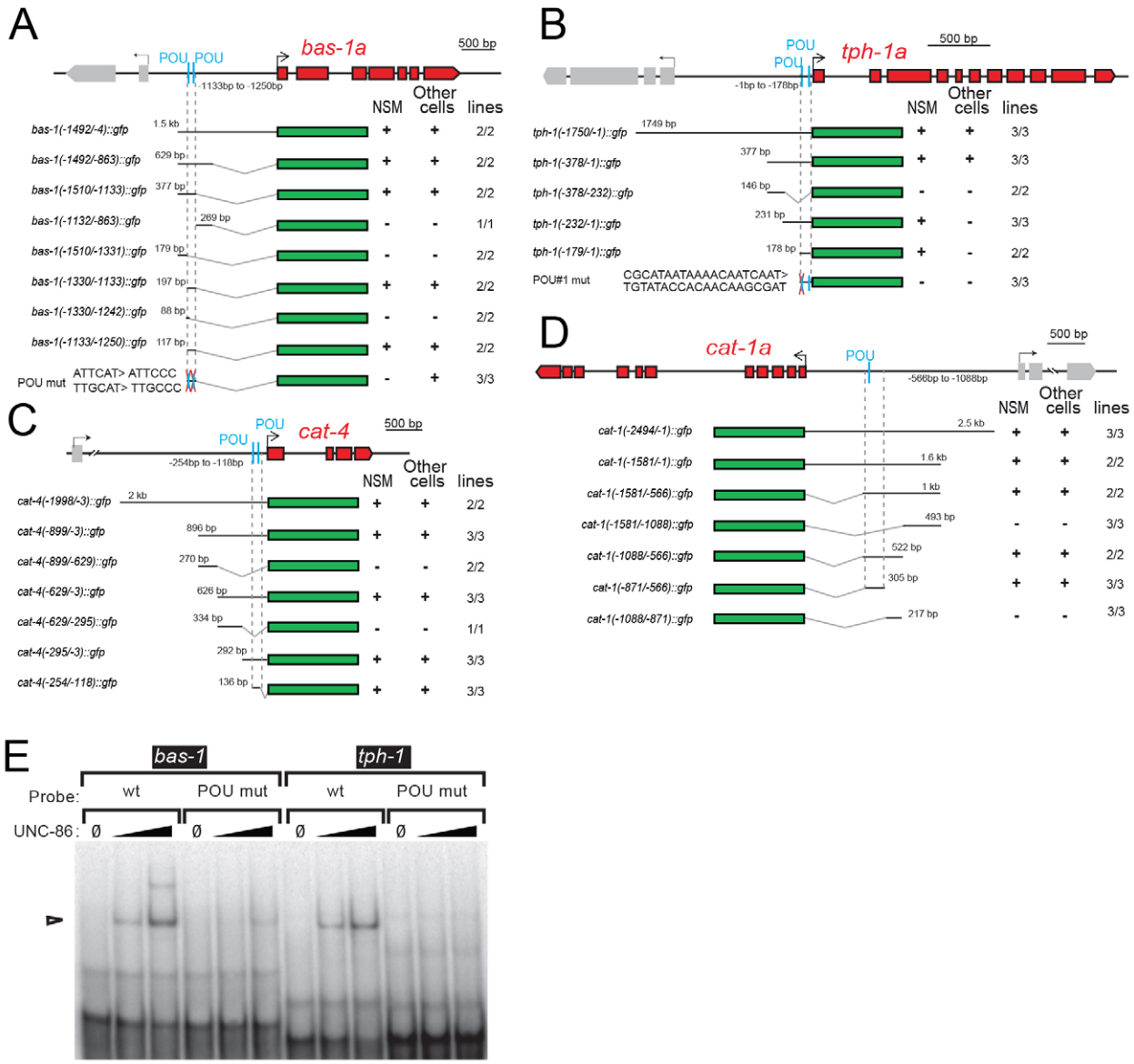


Fig. 6. *Cis*-regulatory analysis of NSM identity specification. (A-D) Dissection of the *cis*-regulatory elements of four NSM-expressed serotonin pathway genes. All minimal *cis*-regulatory elements contain predicted POU sites. (E) EMSAs with UNC-86 protein on *bas-1* and *tph-1* regulatory elements. Mutated POU sites are those that also disrupt reporter gene activity when deleted from the gene contexts of *bas-1* and *tph-1* (A). EMSA was performed with 10 nM and 30 nM UNC-86. Arrowhead indicates UNC86-bound DNA probe.

transporter *eat-4* is ectopically expressed in NSMs (supplementary material Fig. S1A).

unc-86 controls terminal differentiation of the cholinergic IL2, URA and URB sensory, motor and interneurons

Apart from our description of *unc-86* terminal selector function in the serotonergic NSM neurons, *unc-86* had previously been described to broadly affect the terminal differentiation program of other serotonergic (Sze et al., 2002) as well as glutamatergic (Duggan et al., 1998; Serrano-Saiz et al., 2013) neurons. We asked whether *unc-86* might affect the terminal differentiation program of neurons that use yet another neurotransmitter system. We turned to the six IL2 sensory neurons that are involved in nictation behavior (Lee et al., 2012). The IL2 neurons express *unc-86* throughout their lifetime and have been inferred to be cholinergic (Lee et al., 2012). We corroborated the cholinergic identity of the IL2s by finding that reporter fusions to the *unc-17/cha-1* locus and to the choline reuptake transporter *cho-1* are expressed in IL2 neurons (Fig. 8A). The expression of these two key markers of cholinergic identity is eliminated in *unc-86* mutants (Fig. 8A). Expression of the nicotinic acetylcholine receptor subunit *des-2* is also lost in the IL2 neurons of *unc-86* mutants (Treinin et al., 1998).

In addition to these cholinergic markers, we examined the expression of other genes previously shown to be expressed in IL2 neurons, namely the *unc-5* netrin receptor, the guanylyl cyclase *gcy-19*, the kinesin *klp-6* and the Notch ligand *lag-2* (which is expressed in IL2 neurons at the dauer stage) (Leung-Hagesteijn et al., 1992; Ortiz et al., 2006; Ouellet et al., 2008; Peden and Barr, 2005). The expression of all of these terminal markers of IL2 identity is eliminated in IL2 neurons of *unc-86* mutants (Fig. 8A). IL2 neurons also fail to take up dye in *unc-86* mutants (Tong and Bürglin, 2010), suggesting morphological defects. The IL2 neurons are nevertheless

Development (2014) doi:10.1242/dev.099721

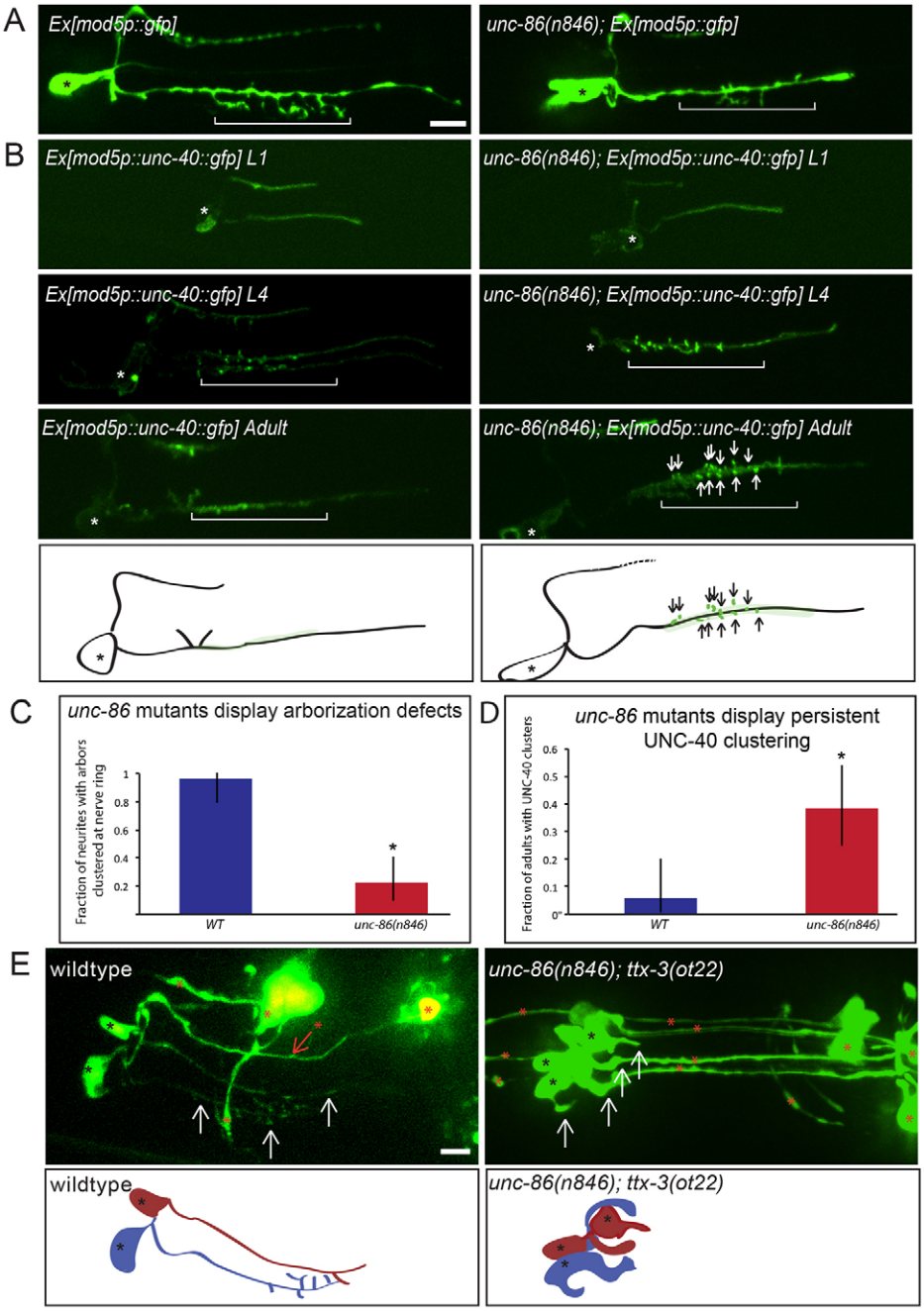


Fig. 7. *unc-86* and *ttx-3* affect NSM morphology. (A) *unc-86(n846)* mutant adults display shorter ventral neurites and fewer and shorter axon arbors. *mod-5p::gfp* (*olaEx1446*) is used to visualize NSM morphology. (B) UNC-40::GFP localization (transgene: *olaEx1448*) remains in a juvenile state in *unc-86* mutant animals. White arrows indicate UNC-40::GFP puncta. (C) Quantification of the *unc-86(n846)* arborization phenotype. Displayed is the fraction of animals with clusters of arbors in the nerve ring region in wild-type and *unc-86(n846)* animals. The difference between wild type and *unc-86* is significant (**P*<0.0001). (D) Quantification of the *unc-86(n846)* UNC-40::GFP localization phenotype. The fraction of animals with multiple, bright UNC-40::GFP puncta in the nerve ring region is displayed. The difference between wild type and *unc-86* is significant (**P*=0.0017). Error bars indicate 95% confidence intervals. (E) *unc-86(n846); ttx-3(ot22)* double mutants display numerous NSM morphology defects, as visualized with *flp-4::gfp (olaEx1485*). In all images, anterior is to the left and ventral down. Asterisks indicate cell bodies (A,B) or additional cell-body-like swellings (E). Brackets denote the nerve ring terminal field where arbors form. White arrows indicate NSM neurites, red arrows and asterisks denote other non-NSM structures. Fisher's *t*-test was used for statistical analysis. Scale bars: 5 µm.

generated in *unc-86* mutants, as assessed by intact expression of the pan-neuronal marker *rab-3* and the pan-sensory marker *osm-6* (50 animals were scored for each marker).

unc-86 is expressed in two additional cholinergic neuron classes in the anterior ganglion besides the IL2 sensory neurons, namely the URA motoneurons [which are synaptically connected to the IL2 neurons (White et al., 1986)] and the URB interneurons. We found that cholinergic identity was also strongly affected in both URA and URB neurons of *unc-86(n846)* loss-of-function mutants (supplementary material Fig. S4).

unc-86 cooperates with the ARID transcription factor *cfi-1* to control IL2 and URA identity

Since none of the previously known co-factors of *unc-86* [*mec-3* for touch neurons (Duggan et al., 1998) and *ttx-3* for NSM neurons (this paper)] is expressed in IL2, URA or URB neurons, *unc-86* is likely to act with another co-factor in IL2 neurons. *cfi-1* is an ARID transcription factor previously shown to be co-expressed with *unc-86* specifically in IL2 and URA neurons (Shaham and Bargmann, 2002). Loss of *cfi-1* results in ectopic expression of identity markers for the CEM neuron in IL2 and URA neurons (Shaham and Bargmann, 2002), which prompted us to investigate whether *cfi-1* might also positively control their cholinergic identity. We find that the cholinergic identity of both IL2 and URA neurons is affected in *cfi-1(ky651)* loss-of-function mutants, albeit not as strongly as in *unc-86* null mutants (Fig. 8A; supplementary material Fig. S4). To investigate whether *unc-86* and *cfi-1* genetically interact, we examined non-additive synergistic interactions of the two genes using a hypomorphic *unc-86* allele, *n848*. Animals carrying this allele show mild IL2 and URA differentiation defects, but in

Development (2014) doi:10.1242/dev.099721

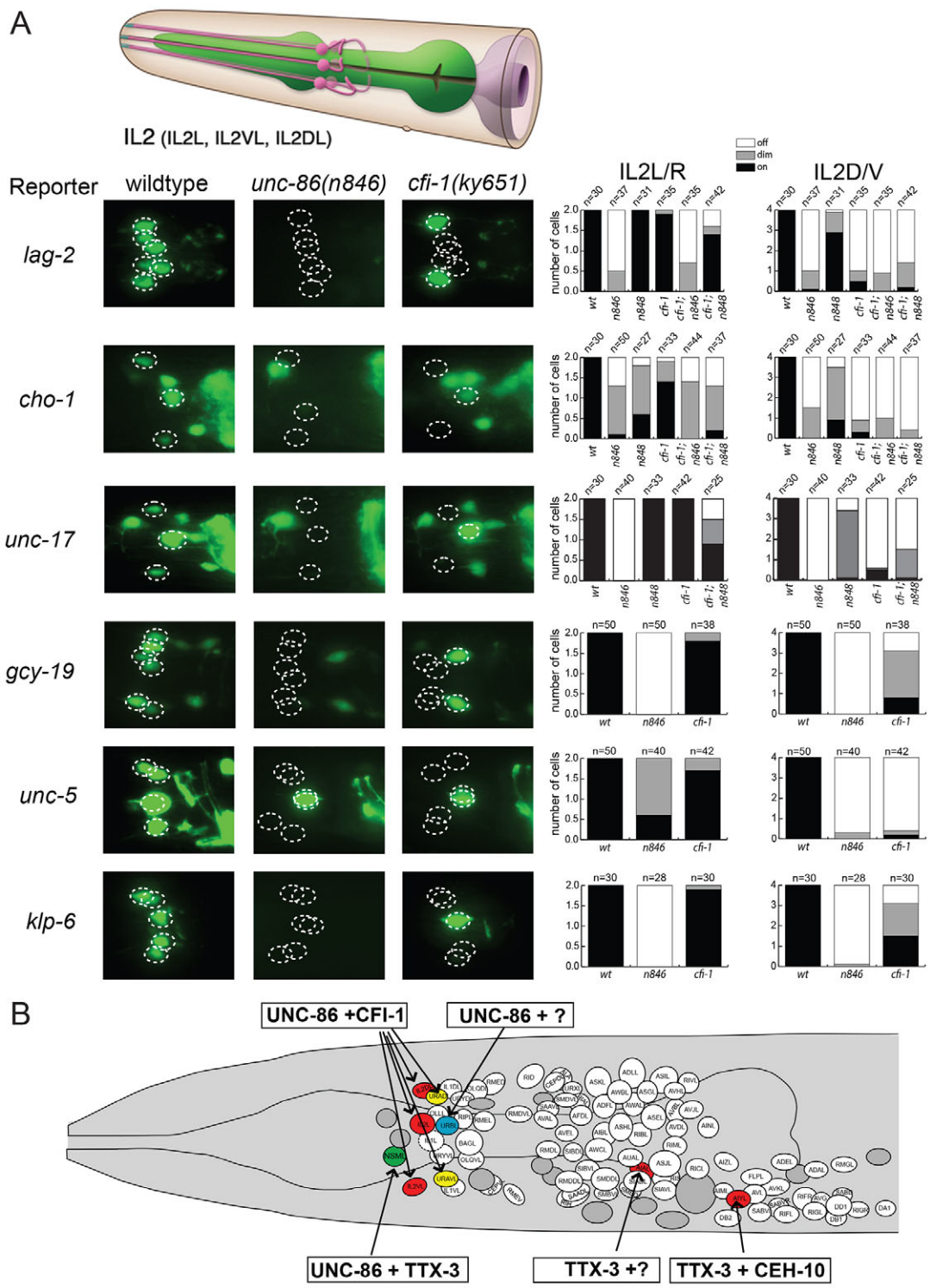


Fig. 8. *unc-86* and *cfi-1* control cholinergic IL2 neuron identity. (A) Animals are late L4 or young adults, with the exception of the *lag-2::gfp* transgenic animals which are dauers. The differential importance of *cfi-1* in the dorsal IL2DL/R and ventral IL2VL/R neurons versus the lateral IL2L/R neurons mirrors morphological differences of the ventral versus lateral neurons, with the lateral neurons having a distinct spectrum of synaptic partners (White et al., 1986). See also supplementary material Fig. S4. IL2 schematic reproduced with permission (Altun et al., 2002-2013). (B) Summary of terminal selector combinatorial codes in head ganglia of *C. elegans*. Colors refer to neurotransmitter identities: green, serotonergic; red, cholinergic; yellow, glutamatergic. Support or blast cells are in gray.

combination with the *cfi-1(ky651)* mutant allele there are strong synergistic, i.e. non-additive, defects in IL2 and URA differentiation (Fig. 8A; supplementary material Fig. S4). We conclude that *unc-86* and *cfi-1* cooperate to control IL2 and URA identity.

DISCUSSION

Two main conclusions can be drawn from the data presented in this paper. First, our data provide general support for the terminal selector concept. Second, our data show that a given transcription

Development (2014) doi:10.1242/dev.099721

factor can operate as a selector of terminal neuron identity in distinct neuronal cell types and that this is achieved through cooperation with distinct co-factors (summarized in Fig. 8B). In other words, individual neuronal cell types use distinct combinatorial codes of terminal selectors, and individual components of the code are reused in distinct combinations in different cell types.

The terminal selector concept was initially proposed based on a relatively small number of *C. elegans* transcription factor mutant phenotypes (Hobert, 2008). In each of these mutant backgrounds, a neuronal cell is born and expresses pan-neuronal features but fails to adopt neuron type-specific identity features. Importantly, terminal differentiation is very broadly affected in terminal selector mutants, such that not only functionally linked features (such as enzymes and transporter in a neurotransmitter synthesis/transport pathway), but also seemingly completely independent differentiation features that have no obvious biochemical connection (e.g. sensory receptors, neuropeptides and ionotropic neurotransmitter receptors) fail to be expressed. That the removal of an individual transcription factor results in such broad defects could not necessarily be assumed since transcriptomic approaches generally show that individual cell types expresses several dozen transcription factors (e.g. Etchberger et al., 2007). This could be interpreted to mean that the identity features of a neuron are regulated in a piecemeal manner, rather than being 'mastered' by a single transcription factor or a small combination thereof (Hobert, 2011). Two major questions raised by the terminal selector concept were how broadly it applies to different cell types in the *C. elegans* nervous system and how it applies to transcription factors expressed in distinct neuron types.

Here, we have shown that the terminal differentiation programs of very distinct neuron types – a cholinergic interneuron (AIA), a serotonergic sensory/motor neuron (NSM) and cholinergic sensory and motor neuron classes (IL2 and URA) – are controlled by distinct combinatorial codes of transcription factors. These factors regulate many distinct identity features of these distinct neuron types, ranging from neuropeptides to neurotransmitter synthesis pathway genes to neurotransmitter receptors and other signaling molecules.

In the case of the cholinergic AIA interneuron, we found that the expression of every tested terminal differentiation marker is affected in *ttx-3* mutants. Since the available AIA marker collection essentially represents a random snapshot of terminal markers that characterize AIA identity, one might extrapolate the regulatory impact of *ttx-3* on each one of these genes to the many hundreds, if not thousands, of genes that are expressed in AIAs, such that *ttx-3* is likely to affect a very large number of them. The estimated very broad effect of *ttx-3* on AIA identity is consistent with what we observed for the cholinergic AIY interneuron, in which *ttx-3* mutation also affects the expression of all known identity features (Altun-Gultekin et al., 2001; Wenick and Hobert, 2004). Even though both neuron types have similar morphologies, are cholinergic, and are directly postsynaptic to various sensory neurons, AIY and AIA have different functions (Hobert et al., 1997; Shinkai et al., 2011; Tomioka et al., 2006), connect to a different spectrum of synaptic partners (White et al., 1986) and express distinct gene batteries. Yet, in both cases, *ttx-3* very broadly affects the differentiation of each neuron type.

The distinct target gene specificities of *ttx-3* in AIA and AIY neurons can be explained by neuron type-specific co-factors and by *ttx-3* acting through distinct *cis*-regulatory motifs. AIY-expressed genes display a characteristic *cis*-regulatory signature that is recognized by a combination of the TTX-3 and CEH-10 homeodomain proteins (Wenick and Hobert, 2004). As we have shown here, AIA-expressed genes share a distinct *cis*-regulatory signature that is composed of two separate motifs located in close proximity, one a TTX-3 binding site and the other a binding site for a presumptive TTX-3 co-factor. This is analogous to the situation in the AIY interneuron class, in which TTX-3 and CEH-10 operate through a bipartite motif (the 'AIY motif') composed of a TTX-3 and a CEH-10 binding site (Wenick and Hobert, 2004). Genes that are expressed in both AIY and AIA neurons (e.g. *cho-1*) contain a modular assembly of both the AIY and AIA *cis*-regulatory signature.

Similar to the *ttx-3*-dependent control of the central cholinergic interneurons AIY and AIA, the mouse LIM homeobox gene *Lhx7* is required for the terminal differentiation of cholinergic striatal interneurons (Lopes et al., 2012). As with other terminal selector transcription factors, *Lhx7* function appears to be continuously required to maintain cholinergic identity. Co-factors that operate together with *Lhx7* are currently not known. *Lhx7* is expressed in many other neurons in the CNS. It will be interesting to determine whether *Lhx7* also operates as a terminal selector in these other neuron types.

ttx-3 activity is not restricted to cholinergic neurons. We find that *ttx-3* is also a key regulator of serotonergic neuron identity. The activity of *ttx-3* in the serotonergic NSM neuron class is, however, distinct from that of AIA and AIY. Whereas the expression of several NSM-expressed effector genes is completely eliminated in *ttx-3* mutants, the expression of some effector genes is only partially affected or not affected at all. In cases in which only partial or no effect was observed, joint removal of another homeobox gene, *unc-86*, resulted in much stronger or complete loss of effector gene expression. Vice versa, the expression of effector genes that are unaffected in expression in *unc-86* mutants is lost in either *ttx-3* mutants or in the *ttx-3; unc-86* double mutant. Taken together, elimination of both of the POU/LIM homeobox genes *unc-86* and *ttx-3* has profound effects on NSM identity, paralleling the profound effect that another POU/LIM homeobox combination (*unc-86* and *mec-3*) has on touch neuron differentiation (Duggan et al., 1998). How *unc-86* and *ttx-3* interact to control NSM differentiation is currently unclear. Both genes are continuously expressed in NSM neurons, but do not regulate the expression of each other. Based on the synergistic nature of the effect of joint *ttx-3* and *unc-86* removal on the expression of some target genes (no or limited effect in single mutants, complete loss in double mutant), we propose that both transcription factors act jointly on common target gene promoters. For some target genes, the loss of one regulatory factor can be completely or partly compensated for by the other regulatory factor; in other cases, such compensation is not possible. *unc-86* and *ttx-3* might therefore not always act in a strict cooperative sense, but rather act independently on target gene promoters. There is already a notable precedent for such a mechanism, as we recently found that a combination of three different transcription factors controls dopaminergic neuron identity. For some target genes, individual transcription factor mutants display very limited effects, but double mutants strongly affect target gene expression (Doitsidou et al., 2013). In the case of NSM, we cannot however rule out the possibility that some genes are exclusively regulated by *unc-86* whereas others are exclusively regulated by *ttx-3*.

Apart from demonstrating *ttx-3* terminal selector function in distinct neuron types, we have also shown here that the POU homeobox gene *unc-86* can similarly act as a terminal selector in distinct neuron types. A role of *unc-86* in the differentiation of serotonergic and glutamatergic touch neurons has been described previously (Desai et al., 1988; Duggan et al., 1998; Sze et al., 2002; Serrano-Saiz et al., 2013). We show here that *unc-86* also controls the terminal differentiation programs of three distinct cholinergic

neuron types. Two of these cholinergic neuron types are synaptically connected and form a simple sensory-to-motor circuit (White et al., 1986). The role of *unc-86* in controlling cholinergic IL2 sensory neuron specification is reminiscent of, and might even be homologous to, the function of the POU homeobox gene *acj6* in controlling expression of the cholinergic gene locus in *Drosophila* olfactory neurons (Lee and Salvaterra, 2002). The ARID-type transcription factor *cfi-1* cooperates with *unc-86* to control the cholinergic identity of IL2 and URA neurons. Although neuronal differentiation functions have been reported for the *cfi-1* homolog *dead ringer* (*retained* – FlyBase) in *Drosophila* (Ditch et al., 2005), the functions of vertebrate orthologs (*Arid3* genes) in the nervous system remain to be explored.

MATERIALS AND METHODS

Strains and transgenes

For a list of strains and transgenes and notes on their generation see supplementary material Table S2.

Serotonin antibody staining

Young adult animals were fixed in 4% paraformaldehyde overnight and then treated with 5% β-mercaptoethanol overnight followed by 1000 units/ml collagenase (Sigma-Aldrich) treatment. Rabbit anti-serotonin whole serum (Sigma-Aldrich, S5545) was used at 1:100 dilution. Worms were then washed and incubated with Alexa Fluor 555 donkey anti-rabbit IgG (1:1000; Life Technologies, A-31571).

Cis-regulatory analysis

DNA sequences were subcloned into pPD95.75 expression vector (Addgene). For some smaller constructs, PCR products were directly amplified from subcloned constructs that have the same 3′ end of the promoter sequences. DNAs for injection were PCR amplified to eliminate vector backbone, gel purified and then injected as complex arrays (10 ng/µl) with digested *rol-6(d)* (3 ng/µl) as injection marker, or plasmid mix was directly injected [50 ng/µl together with 100 ng/µl *rol-6(d)*].

Gel shift analysis

Full-length *unc-86* cDNA was cloned into the pET-21b His tag expression vector (EMD Millipore) and transformed into BL21(DE3) pLysS bacteria (Novagen). Protein expression was induced using 1 mM IPTG for 4 hours at 37°C and batch purified using Ni-NTA resin (Qiagen) under denaturing conditions as described (Wenick and Hobert, 2004). TTX-3 was purified and electrophoretic mobility shift assays (EMSAs) were performed as described (Wenick and Hobert, 2004). Probe sequences are listed in supplementary material Table S3.

Acknowledgements

We thank Q. Chen and B. Alarcon for expert assistance in generating transgenic strains, V. Reinke for providing the *ttx-3* fosmid reporter, E. Serrano for *eat-4* reporters, members of the worm community for providing reporter genes and members of the O.H. laboratory for comments on the manuscript.

Competing interests

The authors declare no competing financial interests.

Author contributions

F.Z. and O.H. initiated the study. A.B. and P.G. performed the analysis of the IL2 neurons; A.B. performed analysis of the URA and URB neurons; C.L.-F., M.M. and N.F. undertook the mutational analysis of the serotonergic pathway promoters; J.C.N. and D.A.C.-R. performed and supervised the morphological analysis of the NSM neurons; N.A. and R.S.M. performed and supervised the gel-shift analyses. F.Z. performed all other experiments. O.H. wrote the paper.

Funding

This work was funded by the National Institutes of Health [R01NS039996-05 and R01NS050266-03 to O.H.; R01NS076558 to D.A.C.-R.; and R01NS070644 to R.S.M.]; a March of Dimes Foundation Grant (to D.A.C.-R.); the Spanish Government [SAF2011-26273 to N.F.]; a Marie Curie Career Integration Grant (to N.F.); a VAL i+d Fellowship from Generalitat Valenciana (to C.L.-F.); and a European Research Council Starting Grant (to N.F.). N.F. is a National Alliance for Research on Schizophrenia and Depression (NARSAD) Young Investigator. O.H. is an Investigator of the Howard Hughes Medical Institute. Deposited in PMC for release after 6 months.

Supplementary material

Supplementary material available online at
http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.099721/-/DC1

References

Albertson, D. G. and Thomson, J. N. (1976). The pharynx of Caenorhabditis elegans. *Philos. Trans. R. Soc. B* **275**, 299-325.

Altun, Z. F., Herndon, L. A., Crocker, C., Lints, R. and Hall, D. H. (**eds**) (2002-2013). WormAtlas, http://www.wormatlas.org.

Altun-Gultekin, Z., Andachi, Y., Tsalik, E. L., Pilgrim, D., Kohara, Y. and Hobert, O. (2001). A regulatory cascade of three homeobox genes, ceh-10, ttx-3 and ceh-23, controls cell fate specification of a defined interneuron class in C. elegans. *Development* **128**, 1951-1969.

Aspöck, G., Ruvkun, G. and Bürglin, T. R. (2003). The Caenorhabditis elegans ems class homeobox gene ceh-2 is required for M3 pharynx motoneuron function. *Development* **130**, 3369-3378.

Axäng, C., Rauthan, M., Hall, D. H. and Pilon, M. (2008). Developmental genetics of the C. elegans pharyngeal neurons NSML and NSMR. *BMC Dev. Biol.* **8**, 38.

Baumeister, R., Liu, Y. and Ruvkun, G. (1996). Lineage-specific regulators couple cell lineage asymmetry to the transcription of the Caenorhabditis elegans POU gene unc-86 during neurogenesis. *Genes Dev.* **10**, 1395-1410.

Berger, M. F., Badis, G., Gehrke, A. R., Talukder, S., Philippakis, A. A., Peña-Castillo, L., Alleyne, T. M., Mnaimneh, S., Botvinnik, O. B., Chan, E. T. et al. (2008). Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. *Cell* **133**, 1266-1276.

Bertrand, V. and Hobert, O. (2009). Linking asymmetric cell division to the terminal differentiation program of postmitotic neurons in C. elegans. *Dev. Cell* **16**, 563-575.

Clark, S. G. and Chiu, C. (2003). C. elegans ZAG-1, a Zn-finger-homeodomain protein, regulates axonal development and neuronal differentiation. *Development* **130**, 3781-3794.

Deneris, E. S. and Wyler, S. C. (2012). Serotonergic transcriptional networks and potential importance to mental health. *Nat. Neurosci.* **15**, 519-527.

Desai, C., Garriga, G., McIntire, S. L. and Horvitz, H. R. (1988). A genetic pathway for the development of the Caenorhabditis elegans HSN motor neurons. *Nature* **336**, 638-646.

Ditch, L. M., Shirangi, T., Pitman, J. L., Latham, K. L., Finley, K. D., Edeen, P. T., Taylor, B. J. and McKeown, M. (2005). Drosophila retained/dead ringer is necessary for neuronal pathfinding, female receptivity and repression of fruitless independent male courtship behaviors. *Development* **132**, 155-164.

Doitsidou, M., Flames, N., Topalidou, I., Abe, N., Felton, T., Remesal, L., Popovitchenko, T., Mann, R., Chalfie, M. and Hobert, O. (2013). A combinatorial regulatory signature controls terminal differentiation of the dopaminergic nervous system in C. elegans. *Genes Dev.* **27**, 1391-1405.

Duggan, A., Ma, C. and Chalfie, M. (1998). Regulation of touch receptor differentiation by the Caenorhabditis elegans mec-3 and unc-86 genes. *Development* **125**, 4107-4119.

Etchberger, J. F., Lorch, A., Sleumer, M. C., Zapf, R., Jones, S. J., Marra, M. A., Holt, R. A., Moerman, D. G. and Hobert, O. (2007). The molecular signature and cis-regulatory architecture of a C. elegans gustatory neuron. *Genes Dev.* **21**, 1653-1674.

Finney, M. and Ruvkun, G. (1990). The unc-86 gene product couples cell lineage and cell identity in C. elegans. *Cell* **63**, 895-905.

Flames, N. and Hobert, O. (2009). Gene regulatory logic of dopamine neuron differentiation. *Nature* **458**, 885-889.

Gray, P. A., Fu, H., Luo, P., Zhao, Q., Yu, J., Ferrari, A., Tenzen, T., Yuk, D. I., Tsung, E. F., Cai, Z. et al. (2004). Mouse brain organization revealed through direct genome-scale TF expression analysis. *Science* **306**, 2255-2257.

Greer, E. R., Perez, C. L., Van Gilst, M. R., Lee, B. H. and Ashrafi, K. (2008). Neural and molecular dissection of a C. elegans sensory circuit that regulates fat and feeding. *Cell Metab.* **8**, 118-131.

Harris, G., Korchnak, A., Summers, P., Hapiak, V., Law, W. J., Stein, A. M., Komuniecki, P. and Komuniecki, R. (2011). Dissecting the serotonergic food signal stimulating sensory-mediated aversive behavior in C. elegans. *PLoS ONE* **6**, e21897.

Hirota, J. and Mombaerts, P. (2004). The LIM-homeodomain protein Lhx2 is required for complete development of mouse olfactory sensory neurons. *Proc. Natl. Acad. Sci. USA* **101**, 8751-8755.

Hobert, O. (2008). Regulatory logic of neuronal diversity: terminal selector genes and selector motifs. *Proc. Natl. Acad. Sci. USA* **105**, 20067-20071.

Hobert, O. (2011). Regulation of terminal differentiation programs in the nervous system. *Annu. Rev. Cell Dev. Biol.* **27**, 681-696.

Hobert, O. and Ruvkun, G. (1998). A common theme for LIM homeobox gene function across phylogeny? *Biol. Bull.* **195**, 377-380.

Hobert, O. and Westphal, H. (2000). Functions of LIM-homeobox genes. *Trends Genet.* **16**, 75-83.

Hobert, O., Mori, I., Yamashita, Y., Honda, H., Ohshima, Y., Liu, Y. and Ruvkun, G. (1997). Regulation of interneuron function in the C. elegans thermoregulatory pathway by the ttx-3 LIM homeobox gene. *Neuron* **19**, 345-357.
Horvitz, H. R., Chalfie, M., Trent, C., Sulston, J. E. and Evans, P. D. (1982). Serotonin and octopamine in the nematode Caenorhabditis elegans. *Science* **216**, 1012-1014.
Jafari, G., Xie, Y., Kullyev, A., Liang, B. and Sze, J. Y. (2011). Regulation of extrasynaptic 5-HT by serotonin reuptake transporter function in 5-HT-absorbing neurons underscores adaptation behavior in Caenorhabditis elegans. *J. Neurosci.* **31**, 8948-8957.
Kolterud, A., Alenius, M., Carlsson, L. and Bohm, S. (2004). The Lim homeobox gene Lhx2 is required for olfactory sensory neuron identity. *Development* **131**, 5319-5326.
Lee, M. H. and Salvaterra, P. M. (2002). Abnormal chemosensory jump 6 is a positive transcriptional regulator of the cholinergic gene locus in Drosophila olfactory neurons. *J. Neurosci.* **22**, 5291-5299.
Lee, R. Y., Sawin, E. R., Chalfie, M., Horvitz, H. R. and Avery, L. (1999). EAT-4, a homolog of a mammalian sodium-dependent inorganic phosphate cotransporter, is necessary for glutamatergic neurotransmission in caenorhabditis elegans. *J. Neurosci.* **19**, 159-167.
Lee, H., Choi, M. K., Lee, D., Kim, H. S., Hwang, H., Kim, H., Park, S., Paik, Y. K. and Lee, J. (2012). Nictation, a dispersal behavior of the nematode Caenorhabditis elegans, is regulated by IL2 neurons. *Nat. Neurosci.* **15**, 107-112.
Lein, E. S., Hawrylycz, M. J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A. F., Boguski, M. S., Brockway, K. S., Byrnes, E. J. et al. (2007). Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168-176.
Leung-Hagesteijn, C., Spence, A. M., Stern, B. D., Zhou, Y., Su, M. W., Hedgecock, E. M. and Culotti, J. G. (1992). UNC-5, a transmembrane protein with immunoglobulin and thrombospondin type 1 domains, guides cell and pioneer axon migrations in C. elegans. *Cell* **71**, 289-299.
Liu, C., Maejima, T., Wyler, S. C., Casadesus, G., Herlitze, S. and Deneris, E. S. (2010). Pet-1 is required across different stages of life to regulate serotonergic function. *Nat. Neurosci.* **13**, 1190-1198.
Lopes, R., Verhey van Wijk, N., Neves, G. and Pachnis, V. (2012). Transcription factor LIM homeobox 7 (Lhx7) maintains subtype identity of cholinergic interneurons in the mammalian striatum. *Proc. Natl. Acad. Sci. USA* **109**, 3119-3124.
Mangale, V. S., Hirokawa, K. E., Satyaki, P. R., Gokulchandran, N., Chikbire, S., Subramanian, L., Shetty, A. S., Martynoga, B., Paul, J., Mai, M. V. et al. (2008). Lhx2 selector activity specifies cortical identity and suppresses hippocampal organizer fate. *Science* **319**, 304-309.
Moreno, N., Bachy, I., Rétaux, S. and González, A. (2005). LIM-homeodomain genes as territory markers in the brainstem of adult and developing Xenopus laevis. *J. Comp. Neurol.* **485**, 240-254.
Nathoo, A. N., Moeller, R. A., Westlund, B. A. and Hart, A. C. (2001). Identification of neuropeptide-like protein gene families in Caenorhabditis elegans and other species. *Proc. Natl. Acad. Sci. USA* **98**, 14000-14005.
Nelson, J. C. and Colón-Ramos, D. A. (2013). Serotonergic neurosecretory synapse targeting is controlled by netrin-releasing guidepost neurons in Caenorhabditis elegans. *J. Neurosci.* **33**, 1366-1376.
Ortiz, C. O., Etchberger, J. F., Posy, S. L., Frøkjaer-Jensen, C., Lockery, S., Honig, B. and Hobert, O. (2006). Searching for neuronal left/right asymmetry: genomewide analysis of nematode receptor-type guanylyl cyclases. *Genetics* **173**, 131-149.
Ouellet, J., Li, S. and Roy, R. (2008). Notch signalling is required for both dauer maintenance and recovery in C. elegans. *Development* **135**, 2583-2592.
Peden, E. M. and Barr, M. M. (2005). The KLP-6 kinesin is required for male mating behaviors and polycystin localization in Caenorhabditis elegans. *Curr. Biol.* **15**, 394-404.
Peukert, D., Weber, S., Lumsden, A. and Scholpp, S. (2011). Lhx2 and Lhx9 determine neuronal differentiation and compartition in the caudal forebrain by regulating Wnt signaling. *PLoS Biol.* **9**, e1001218.
Ranganathan, R., Sawin, E. R., Trent, C. and Horvitz, H. R. (2001). Mutations in the Caenorhabditis elegans serotonin reuptake transporter MOD-5 reveal serotonin-dependent and -independent activities of fluoxetine. *J. Neurosci.* **21**, 5871-5884.
Rhee, J. M., Gruber, C. A., Brodie, T. B., Trieu, M. and Turner, E. E. (1998). Highly cooperative homodimerization is a conserved property of neural POU proteins. *J. Biol. Chem.* **273**, 34196-34205.
Röhrig, S. (2000). *Modulation of UNC-86 Activity During Caenorhabditis elegans Neurogenesis*. München, Germany: Herbert Utz Verlag.
Serrano-Saiz, E., Poole, R. J., Felton, T., Zhang, F., De La Cruz, E. and Hobert, O. (2013). Modular control of glutamatergic neuronal identity in C. elegans by distinct homeodomain proteins. *Cell* **155**, 659-673.
Shaham, S. and Bargmann, C. I. (2002). Control of neuronal subtype identity by the C. elegans ARID protein CFI-1. *Genes Dev.* **16**, 972-983.
Shinkai, Y., Yamamoto, Y., Fujiwara, M., Tabata, T., Murayama, T., Hirotsu, T., Ikeda, D. D., Tsunozaki, M., Iino, Y., Bargmann, C. I. et al. (2011). Behavioral choice between conflicting alternatives is regulated by a receptor guanylyl cyclase, GCY-28, and a receptor tyrosine kinase, SCD-2, in AIA interneurons of Caenorhabditis elegans. *J. Neurosci.* **31**, 3007-3015.
Simmons, D. K., Pang, K. and Martindale, M. Q. (2012). Lim homeobox genes in the Ctenophore Mnemiopsis leidyi: the evolution of neural cell type specification. *Evodevo* **3**, 2.
Smidt, M. P. and Burbach, J. P. (2009). Terminal differentiation of mesodiencephalic dopaminergic neurons: the role of Nurr1 and Pitx3. *Adv. Exp. Med. Biol.* **651**, 47-57.
Srivastava, M., Larroux, C., Lu, D. R., Mohanty, K., Chapman, J., Degnan, B. M. and Rokhsar, D. S. (2010). Early evolution of the LIM homeobox gene family. *BMC Biol.* **8**, 4.
Sulston, J. E., Schierenberg, E., White, J. G. and Thomson, J. N. (1983). The embryonic cell lineage of the nematode Caenorhabditis elegans. *Dev. Biol.* **100**, 64-119.
Sze, J. Y., Zhang, S., Li, J. and Ruvkun, G. (2002). The C. elegans POU-domain transcription factor UNC-86 regulates the tph-1 tryptophan hydroxylase gene and neurite outgrowth in specific serotonergic neurons. *Development* **129**, 3901-3911.
Tomioka, M., Adachi, T., Suzuki, H., Kunitomo, H., Schafer, W. R. and Iino, Y. (2006). The insulin/PI 3-kinase pathway regulates salt chemotaxis learning in Caenorhabditis elegans. *Neuron* **51**, 613-625.
Tong, Y. G. and Bürglin, T. R. (2010). Conditions for dye-filling of sensory neurons in Caenorhabditis elegans. *J. Neurosci. Methods* **188**, 58-61.
Treinin, M., Gillo, B., Liebman, L. and Chalfie, M. (1998). Two functionally dependent acetylcholine subunits are encoded in a single Caenorhabditis elegans operon. *Proc. Natl. Acad. Sci. USA* **95**, 15492-15495.
Tursun, B., Cochella, L., Carrera, I. and Hobert, O. (2009). A toolkit and robust pipeline for the generation of fosmid-based reporter genes in C. elegans. *PLoS ONE* **4**, e4625.
Wenick, A. S. and Hobert, O. (2004). Genomic cis-regulatory architecture and trans-acting regulators of a single interneuron-specific gene battery in C. elegans. *Dev. Cell* **6**, 757-770.
White, J. G., Southgate, E., Thomson, J. N. and Brenner, S. (1986). The structure of the nervous system of the nematode Caenorhabditis elegans. *Philos. Trans. R. Soc. B* **314**, 1-340.
Zetterström, R. H., Williams, R., Perlmann, T. and Olson, L. (1996). Cellular expression of the immediate early transcription factors Nurr1 and NGFI-B suggests a gene regulatory role in several brain regions including the nigrostriatal dopamine system. *Brain Res. Mol. Brain Res.* **41**, 111-120.